GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 31, 2003, 14:07:03 ; Search time 74 Seconds
(without alignments)
496.989 Million cell updates/sec

Title: US-10-092-404-1
Perfect score: 1522
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1522	100.0	276	20	AAW94295	Wild-type HFE poly
2	1522	100.0	348	18	AAW36499	Hereditary haemoch
3	1522	100.0	348	21	AAB19149	A human histocompa
4	1522	100.0	348	22	AAB36869	Human hereditary h
5	1517	99.7	438	23	AAU80035	Beta 2 microglobul
6	1513	99.4	276	20	AAW94296	HFE mutant (H63D-H
7	1513	99.4	348	22	AAB36871	Human hereditary h
8	1511	99.3	348	22	AAB36870	Human hereditary h
9	1502	98.7	276	20	AAW94297	HFE mutant (H111A/
10	1502	98.7	348	22	AAB36872	Human hereditary h
11	522	34.3	361	22	AAB36873	Rabbit leukocyte a
12	513	33.7	365	22	AAB36874	MHC class I protei
13	505	33.2	274	21	AAY68275	Human leukocyte an
14	505	33.2	274	21	AAY52929	HLA-A2/A28 family
15	505	33.2	274	22	AAB58690	HLA-A2/A28 protein
16	505	33.2	280	22	AAU10225	Human leukocyte an
17	505	33.2	415	22	AAU10224	Human partial beta
18	504	33.1	365	21	AAY68265	Human leukocyte an
19	504	33.1	365	21	AAY52919	HLA-A2/A28 family
20	504	33.1	365	22	AAB58680	HLA-A2/A28 protein
21	504	33.1	368	22	AAM24017	Human EST encoded
22	503	33.0	274	21	AAY68276	Human leukocyte an
23	503	33.0	274	21	AAY52930	HLA-A2/A28 family
24	503	33.0	274	22	AAB58691	HLA-A2/A28 protein
25	503	33.0	365	21	AAY68268	Human leukocyte an
26	503	33.0	365	21	AAY52922	HLA-A2/A28 family
27	503	33.0	365	22	AAB58683	HLA-A2/A28 protein
28	502	33.0	274	9	AAP80911	Consensus sequence
29	502	33.0	365	21	AAY68267	Human leukocyte an
30	502	33.0	365	21	AAY52921	HLA-A2/A28 family
31	502	33.0	365	22	AAB58682	HLA-A2/A28 protein
32	501	32.9	274	21	AAY68274	Human leukocyte an
33	501	32.9	274	21	AAY52928	HLA-A2/A28 family
34	501	32.9	274	22	AAB58689	HLA-A2/A28 protein
35	501	32.9	365	21	AAY68266	Human leukocyte an
36	501	32.9	365	21	AAY52920	HLA-A2/A28 family
37	501	32.9	365	22	AAB58681	HLA-A2/A28 protein
38	500	32.9	412	19	AAW68385	Chimeric HLA-A2.1/
39	499	32.8	274	21	AAY68273	Human leukocyte an
40	499	32.8	274	21	AAY52927	HLA-A2/A28 family
41	499	32.8	274	22	AAB58688	HLA-A2/A28 protein
42	497	32.7	365	21	AAY68270	Human leukocyte an
43	497	32.7	365	21	AAY68272	Human leukocyte an
44	497	32.7	365	21	AAY52924	HLA-A2/A28 family
45	497	32.7	365	21	AAY52926	HLA-A2/A28 family

ALIGNMENTS

RESULT 1
AAW94295
ID AAW94295 standard; peptide; 276 AA.
XX
AC AAW94295;
XX
DT 27-APR-1999 (first entry)
XX
DE Wild-type HFE polypeptide sequence.
XX
KW HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW transfusion; protein replacement therapy; hereditary hemochromatosis;
KW transferrin receptor; iron deficiency; anemia.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "indicated in the sequence listing as Arg"
XX
PN WO9856814-A1.
XX
PD 17-DEC-1998.
XX
PF 12-JUN-1998; 98WO-US12436.
XX
PR 13-JUN-1997; 97US-0876010.
XX
PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
PA (PROG-) PROGENITOR INC.
XX
PI Bjorkman PJ, Feder JN, Schatzman RC;
XX

```
DR   WPI; 1999-080886/07.
XX
PT   New treatment of an iron overload disease - comprises use of HFE
PT   polypeptides provided in a complex with full length, wild type human
PT   (2m), useful in protein replacement therapy
XX
PS   Claim 1; Page 13; 36pp; English.
XX
CC   The present sequence represents a wild-type HFE polypeptide. The HFE
CC   polypeptides (AAW94295-297)  provided in a complex with full length,
CC   wild type human beta-2-microglobulin (beta2m) form compositions in the
CC   treatment of primary iron overload diseases (e.g. hemochromatosis), or
CC   other iron overload conditions resulting from secondary causes (e.g.
CC   repeated transfusions). Data regarding the structure and function
CC   correlations of HFE polypeptides is useful in designing drugs that
CC   modulate the HFE gene and HFE activity. The polypeptides are also useful
CC   in protein replacement therapy for individuals possessing a defective
CC   HFE gene (e.g. Hereditary hemochromatosis). (Ant)agonists of the
CC   polypeptides are also useful in treating primary and secondary iron
CC   overload diseases. The modulators of the transferrin receptor are useful
CC   in treating iron deficiency conditions such as anemia, and in modulating
CC   the amount of iron transported into a cell. The HFE polypeptides provide
CC   a molecular basis for the relationship between HFE and iron metabolism,
CC   which enables treatment of iron overload and deficiency diseases.
XX
SQ   Sequence   276 AA;

  Query Match             100.0%;  Score 1522;  DB 20;  Length 276;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-135;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||||||||||||||||||||||||||||||||||||
Db     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 2
AAW36499
ID   AAW36499 standard; Protein; 348 AA.
XX
AC   AAW36499;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene product.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 63
FT                   /note= "substituted by Asp in 24s2 mutant"
FT   Misc-difference 65
FT                   /note= "substituted by Cys in 24d7 variant"
FT   Misc-difference 282
FT                   /note= "substituted by Tyr in 24d1 mutant"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
XX
PA   (MERC-) MERCATOR GENETICS INC.
XX
PI   Drayna DT,  Feder JN,  Gnirke A,  Ruddy D,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1997-512743/47.
DR   N-PSDB; AAT96690-91.
XX
PT   Hereditary haemochromatosis gene and variants - useful for diagnosis
PT   and treatment of hereditary haemochromatosis disease
XX
PS   Disclosure; Fig 4; 115pp; English.
XX
CC   This polypeptide is the expression product of a novel human gene
CC   (see AAT96690) whose mutated form is associated with hereditary
CC   haemochromatosis (HH).  A single mutation (24d1) in the HH gene
CC   appears responsible for the majority of HH disease.  This comprises
CC   a G to A substitution that is present in 86% of affected
CC   chromosomes and in 4% of unaffected chromosomes.  It results in a
CC   Cys to Tyr substitution in the encoded protein at a critical
CC   disulphide bridge important for secondary structure.  The following
CC   are claimed: the 10825 bp genomic DNA sequence (I), a 1437 bp cDNA
CC   sequence (Ia) (see AAT96691) and their 24d1, 24d2 and 24d7 variants;
CC   a cloning or expression vector; host cells; a peptide product
CC   chosen from the HH gene product, its variants (24d1, 24d2 and
CC   24d7), or a peptide of at least 56 amino acid residues of these; an
CC   antibody produced using the peptide as an immunogen; a method to
CC   determine the presence or absence of the common HH gene mutation;
CC   an animal model for the HH disease; metal chelation agents, T-cell
CC   differentiation factors and therapeutic agents for the mitigation
CC   of injury due to oxidative process in vivo or mitigation of iron
CC   overload; a method for screening potential therapeutic agents for
CC   activity in connection with HH disease; an antisense oligonucleotide
CC   directed against a transcriptional product of a nucleic acid
CC   sequence as above; and oligonucleotides or pairs of oligonucleotides
CC   covering a range of nucleotides from (I), (Ia) or their variants,
CC   useful for detecting a polymorphism in the HH gene.  The invention
CC   also relates to methods for screening for HH homozygotes, to HH
CC   diagnosis, prenatal screening and diagnosis, and therapies of HH
CC   disease, including gene therapy, protein- and antibody-based
CC   therapeutics, and small molecule therapeutics.
XX
SQ   Sequence   348 AA;

  Query Match             100.0%;  Score 1522;  DB 18;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-135;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy    241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
          ||||||||||||||||||||||||||||||||||||
Db    263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 3
AAB19149
ID   AAB19149 standard; Protein; 348 AA.
XX
AC   AAB19149;
XX
DT   19-FEB-2001  (first entry)
XX
DE   A human histocompatibility iron loading (HFE) protein.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT                   /note= "signal peptide"
FT   Misc-difference 63
FT                   /note= "when nucleotide 187 is mutated to G, then
FT                         this residue is Asp"
FT   Misc-difference 65
FT                   /note= "when nucleotide 193 is mutated to T, then
FT                         this residue is Cys"
FT   Domain          80..108
FT                   /note= "alphal domain"
FT   Misc-difference 93
FT                   /note= "when nucleotide 277 is mutated to C, then
FT                         this residue is Arg"
FT   Misc-difference 105
FT                   /note= "when nucleotide 314 is mutated to C, then
FT                         this residue is Thr"
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
DR   N-PSDB; AAA96769.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Disclosure; Page 3; 55pp; English.
XX
CC   The present sequence represents a human histocompatibility iron loading
CC   (HFE) protein. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence   348 AA;

  Query Match            100.0%;  Score 1522;  DB 21;  Length 348;
  Best Local Similarity  100.0%;  Pred. No. 4.8e-135;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy     61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy    121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy    181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy    241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
          ||||||||||||||||||||||||||||||||||||
Db    263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 4
AAB36869
ID   AAB36869 standard; Protein; 348 AA.
XX
AC   AAB36869;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK,
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   N-PSDB; AAC68425.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Claim 1; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
```

```
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   348 AA;

  Query Match            100.0%;  Score 1522;  DB 22;  Length 348;
  Best Local Similarity  100.0%;  Pred. No. 4.8e-135;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy     61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy    121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy    181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy    241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
          |||||||||||||||||||||||||||||||||||
Db    263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 5
AAU80035
ID   AAU80035 standard; Protein; 438 AA.
XX
AC   AAU80035;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Beta 2 microglobulin (beta2M)/HFE monochain.
XX
KW   Human; beta 2 microglobulin; beta2M/HFE monochain; HFE; ischaemia;
KW   iron absorption regulator; intracellular iron absorption; lung injury;
KW   haemochromatosis; transfusion; thalassaemia; haemolytic anaemia;
KW   chronic infection; transferrin receptor; TfR; brain tumour; cancer;
KW   oxidative stress disorder; tissue damage; vascular disease;
KW   inflammation; atherosclerosis; autoimmune disease;
KW   inflammatory condition.
XX
OS   Homo sapiens.
XX
PN   WO200224929-A2.
XX
PD   28-MAR-2002.
XX
PF   24-SEP-2001; 2001WO-US29873.
XX
PR   22-SEP-2000; 2000US-234843P.
XX
PA   (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
PA   (MCIN/) MCINNIS P.
XX
PI   Ehrlich R,  Rotem-Yehudar R,  Laham N;
XX
DR   WPI; 2002-383192/41.
DR   N-PSDB; ABK49917.
XX
PT   Soluble beta 2 microglobulin/HFE monochain useful for treating
PT   iron-overload conditions e.g. thalassemia and chronic infections,
PT   comprises human beta 2 microglobulin linked to alpha domains of HFE by
PT   a linker peptide  -
XX
PS   Example 2; Fig 2; 77pp; English.
XX
CC   The invention relates to a soluble polypeptide (I) of beta 2
CC   microglobulin (beta2m)/HFE monochain comprising human beta2m (or its
CC   analogue or active fragment), linked to alpha1-alpha3 domains of human
CC   HFE (a central regulator of iron absorption; undefined), or its analogue
CC   or active fragment, by a flexible linker peptide, or a functional
CC   derivative or salt of (I). (I) is useful for reducing intracellular iron
CC   absorption in patients having hereditary haemochromatosis, transfusions,
CC   thalassaemias, haemolytic anaemia or chronic infections, and for
CC   delivering a therapeutic to cells that over-express transferrin receptor
CC   (TfR) which are preferably lymphocytes or leukocytes, across the blood-
CC   brain barrier. (I) is further useful for treating brain tumour. (I)
CC   is also useful for treating oxidative stress disorders resulting in
CC   tissue damage e.g. vascular diseases, inflammation, atherosclerosis,
CC   lung injury, ischaemia, etc. A DNA molecule (II) encoding (I) is useful
CC   as a platform for drug delivery of therapeutic use for cancer,
CC   autoimmune diseases and inflammatory conditions. The monochain manifests
CC   specific characteristics advantageous for drug delivery systems. It is a
CC   soluble, stable and fully conformed protein. It binds specifically to
CC   transferrin receptor (TfR) and therefore targets cells that over-express
CC   this receptor. It is continuously internalised by the target cells, thus
CC   enabling efficient drug delivery. It dissociates from the receptor in the
CC   cells, minimising side effects. It negatively regulates iron absorption,
CC   reducing growth of undesired cells and preventing lymphocyte activation.
CC   It is not diluted in the blood as is transferrin. It should not induce an
CC   immune response since it is a self non-polymeric protein and delivery of
CC   drugs via monochain is expected to overcome drug-resistance since it is a
CC   natural TfR-binding protein. The present sequence represents the amino
CC   acid sequence of beta2m/HFE monochain.
XX
SQ   Sequence   438 AA;

  Query Match            99.7%;  Score 1517;  DB 23;  Length 438;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-134;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 194

Qy     61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 254

Qy    121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 314

Qy    181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 374

Qy    241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 275
          |||||||||||||||||||||||||||||||||||
Db    375 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 409


RESULT 6
AAW94296
ID   AAW94296 standard; peptide; 276 AA.
XX
AC   AAW94296;
XX
DT   27-APR-1999  (first entry)
XX
DE   HFE mutant (H63D-HFE) polypeptide sequence.
XX
KW   HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW   transfusion; protein replacement therapy; hereditary hemochromatosis;
KW   transferrin receptor; iron deficiency; anemia; mutant.
XX
OS   Synthetic.
XX
```

```
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "indicated in the sequence listing as Arg"
FT   Misc-difference 41
FT                   /label= H63D
FT                   /note= "wild type His (of the mature protein sequence)
FT                          is replaced by Asp"
XX
PN   WO9856814-A1.
XX
PD   17-DEC-1998.
XX
PF   12-JUN-1998;   98WO-US12436.
XX
PR   13-JUN-1997;   97US-0876010.
XX
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
PA   (PROG-) PROGENITOR INC.
XX
PI   Bjorkman PJ,  Feder JN,  Schatzman RC;
XX
DR   WPI; 1999-080886/07.
XX
PT   New treatment of an iron overload disease - comprises use of HFE
PT   polypeptides provided in a complex with full length, wild type human
PT   (2m), useful in protein replacement therapy
XX
PS   Claim 3; Page 14; 36pp; English.
XX
CC   The present sequence represents a H63D-HFE mutant polypeptide. The HFE
CC   polypeptides (AAW94295-297)  provided in a complex with full length,
CC   wild type human beta-2-microglobulin (beta2m) form compositions in the
CC   treatment of primary iron overload diseases (e.g. hemochromatosis), or
CC   other iron overload conditions resulting from secondary causes (e.g.
CC   repeated transfusions). Data regarding the structure and function
CC   correlations of HFE polypeptides is useful in designing drugs that
CC   modulate the HFE gene and HFE activity. The polypeptides are also useful
CC   in protein replacement therapy for individuals possessing a defective
CC   HFE gene (e.g. Hereditary hemochromatosis). (Ant)agonists of the
CC   polypeptides are also useful in treating primary and secondary iron
CC   overload diseases. The modulators of the transferrin receptor are useful
CC   in treating iron deficiency conditions such as anemia, and in modulating
CC   the amount of iron transported into a cell. The HFE polypeptides provide
CC   a molecular basis for the relationship between HFE and iron metabolism,
CC   which enables treatment of iron overload and deficiency diseases.
XX
SQ   Sequence   276 AA;

  Query Match             99.4%;  Score 1513;  DB 20;  Length 276;
  Best Local Similarity   99.6%;  Pred. No. 2.5e-134;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||||||||||||||||||||||||||||||||||
Db     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
```

```
RESULT 7
AAB36871
ID   AAB36871 standard; Protein; 348 AA.
XX
AC   AAB36871;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d2 mutation protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   N-PSDB; AAC68427.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Claim 3; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   348 AA;

  Query Match             99.4%;  Score 1513;  DB 22;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 3.4e-134;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||||||||||||||||||||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298
```

```
RESULT 8
AAB36870
ID   AAB36870 standard; Protein; 348 AA.
XX
AC   AAB36870;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1 mutation protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   N-PSDB; AAC68426.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Claim 2; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   348 AA;

  Query Match           99.3%;  Score 1511;  DB 22;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 5.2e-134;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||||||||||||||||| ||||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298
```

```
RESULT 9
AAW94297
ID   AAW94297 standard; peptide; 276 AA.
XX
AC   AAW94297;
XX
DT   27-APR-1999  (first entry)
XX
DE   HFE mutant (H111A/H145A-HFE) polypeptide sequence.
XX
KW   HFE; beta-2-microglobulin; beta2m; iron overload; hemochromatosis;
KW   transfusion; protein replacement therapy; hereditary hemochromatosis;
KW   transferrin receptor; iron deficiency; anemia; mutant.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "indicated in the sequence listing as Arg"
FT   Misc-difference 89
FT                   /label= H111A
FT                   /note= "wild type His (of the mature protein sequence)
FT                          is replaced by Ala"
FT   Misc-difference 123
FT                   /label= H145A
FT                   /note= "wild type His (of the mature protein sequence)
FT                          is replaced by Ala"
XX
PN   WO9856814-A1.
XX
PD   17-DEC-1998.
XX
PF   12-JUN-1998;   98WO-US12436.
XX
PR   13-JUN-1997;   97US-0876010.
XX
PA   (CALY ) CALIFORNIA INST OF TECHNOLOGY.
PA   (PROG-) PROGENITOR INC.
XX
PI   Bjorkman PJ,  Feder JN,  Schatzman RC;
XX
DR   WPI; 1999-080886/07.
XX
PT   New treatment of an iron overload disease - comprises use of HFE
PT   polypeptides provided in a complex with full length, wild type human
PT   (2m), useful in protein replacement therapy
XX
PS   Claim 5; Page 15; 36pp; English.
XX
CC   The present sequence represents a H111A/H145A-HFE mutant polypeptide.
CC   The HFE polypeptides (AAW94295-297) provided in a complex with full
CC   length, wild type human beta-2-microglobulin (beta2m) form compositions
CC   in the treatment of primary iron overload diseases (e.g.
CC   hemochromatosis), or other iron overload conditions resulting from
CC   secondary causes (e.g. repeated transfusions). Data regarding the
CC   structure and function correlations of HFE polypeptides is useful in
CC   designing drugs that modulate the HFE gene and HFE activity. The
CC   polypeptides are also useful in protein replacement therapy for
CC   individuals possessing a defective HFE gene (e.g. Hereditary
CC   hemochromatosis). (Ant)agonists of the polypeptides are also useful in
CC   treating primary and secondary iron overload diseases. The modulators of
CC   the transferrin receptor are useful in treating iron deficiency
CC   conditions such as anemia, and in modulating the amount of iron
CC   transported into a cell. The HFE polypeptides provide a molecular basis
CC   for the relationship between HFE and iron metabolism, which enables
CC   treatment of iron overload and deficiency diseases.
XX
SQ   Sequence   276 AA;

  Query Match           98.7%;  Score 1502;  DB 20;  Length 276;
  Best Local Similarity   99.3%;  Pred. No. 2.7e-133;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||||||||||||||||||||||||||||||||||||
Db     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 10
AAB36872
ID   AAB36872 standard; Protein; 348 AA.
XX
AC   AAB36872;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1/2 mutation protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   N-PSDB; AAC68428.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Claim 4; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   348 AA;

  Query Match             98.7%;  Score 1502;  DB 22;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 3.7e-133;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||||||||||||||||| ||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 11
AAB36873
ID   AAB36873 standard; Protein; 361 AA.
XX
AC   AAB36873;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Rabbit leukocyte antigen.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Oryctolagus cuniculus.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 7; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   361 AA;                                                ;

  Query Match             34.3%;  Score 522;  DB 22;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 9.1e-41;
  Matches  111;  Conservative   44;  Mismatches  114;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
```

```
          |||: | :   |   ||   |  :|||||  || :| :  | |:| | ||:  ::  : |
Db      26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db      85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           |::    |   | ||:  |  |: :||     | ::||||||:|   |:: ||:|:  |
Db     145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy     181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   |||| |    |||| || :||  |:: | :| :    :: |  :  | ||
Db     204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || ||||||||:|:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299


RESULT 12
AAB36874
ID   AAB36874 standard; Protein; 365 AA.
XX
AC   AAB36874;
XX
DT   21-FEB-2001  (first entry)
XX
DE   MHC class I protein.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload.
XX
OS   Unidentified.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 7; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence   365 AA;

  Query Match             33.7%;  Score 513;  DB 22;  Length 365;
  Best Local Similarity   39.7%;  Pred. No. 6.5e-40;
  Matches  110;  Conservative   45;  Mismatches  114;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |

Db      26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
           :: :|       ||  |:   :| |:  |||||:: ||::  |      || :| |||
Db      85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTLQMMFGCDVGSDWRFLRGYHQYAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db     145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |:|||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 13
AAY68275
ID   AAY68275 standard; Peptide; 274 AA.
XX
AC   AAY68275;
XX
DT   13-APR-2000  (first entry)
XX
DE   Human leukocyte antigen A2/A28 family related protein SEQ ID NO:107.
XX
KW   MHC class I; major histocompatibility complex; microglobulin; antigen;
KW   immune response; immunisation; AIDS; multiple sclerosis; toxic shock;
KW   cancer; lupus erythematosus; snake bite; cytostatic; antiviral;
KW   immunomodulatory; dermatological; immunosuppressive; antiinflammatory;
KW   neuroprotective.
XX
OS   Homo sapiens.
XX
PN   US6011146-A.
XX
PD   04-JAN-2000.
XX
PF   07-JUN-1995;   95US-0481985.
XX
PR   15-NOV-1991;   91US-0792473.
PR   05-DEC-1991;   91US-0801818.
XX
PA   (INSP ) INST PASTEUR.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
PI   Kourilsky P,  Mottez E,  Abastado J;
XX
DR   WPI; 2000-125951/11.
XX
PT   New recombinant DNA encoding covalently linked form of major
PT   histocompatibility complex Class I determinant, used for immune system
PT   stimulation, e.g. for treating cancer  -
XX
PS   Disclosure; Column 127-128; 88pp; English.
XX
CC   The present invention describes a recombinant DNA molecule (I)
CC   containing a sequence (Ia) that encodes an altered MHC (major
CC   histocompatibility complex ) Class I determinant (II) comprises a
CC   polypeptide with alphal, alpha2, alpha3 and beta2-microglobulin
CC   domains, in which alpha3 and beta2 are covalently linked, thorough C-
CC   and N-termini respectively, via a nucleotide spacer sequence encoding a
CC   polypeptide. (II) includes an antigen-binding site and when (II) and
CC   the antigen are associated they are recognized by a mammalian T cell
CC   receptor (TCR). (I) are used to produce (II) which are used to study
CC   functional interactions between the various MHC domains. They can also
CC   be used to modulate (in vivo or in vitro) the immune system by inducing
CC   an effector response (cytotoxicity, antibody synthesis, phagocytosis)
CC   of immune system cells, typically for treating, or immunising against;
CC   cancer, acquired immune deficiency syndrome, lupus erythematosus,
```

```
CC   multiple sclerosis, toxic shock and snake bite, but also for selective
CC   destruction of autoreactive cells, diagnostically to assay T cell
CC   receptors and to raise specific antibodies (useful for diagnosis,
CC   therapy, studying MHC-associated cellular processes and for affinity
CC   purification). AAZ57558 and AAY68186 to AAY68316 are sequences used in
CC   the exemplification of the present invention.
XX
SQ   Sequence   274 AA;

  Query Match             33.2%;  Score 505;  DB 21;  Length 274;
  Best Local Similarity   39.9%;  Pred. No. 2.5e-39;
  Matches  110;  Conservative   43;  Mismatches  115;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |||:||| ||:      : |
Db       2 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASRRMEPRAPWIEQE-GPEYW 60

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
             :: :|      ||  |:   :| |:  ||||| : ||::  |    .|| :| |||
Db      61 DGETRKVKAHSQTHRVDLSTLRGYYNQSEAGSHTLQRMYGCDVGSDWRFLRGYHQYAYDG 120

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db     121 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQWRAYLEGTCVEWLRRYLENGKETL 179

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     180 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 238

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 275
           ||:| |  : || |:|||||| |:| || :|| : |
Db     239 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLPW 274


RESULT 14
AAY52929
ID   AAY52929 standard; Peptide; 274 AA.
XX
AC   AAY52929;
XX
DT   14-FEB-2000  (first entry)
XX
DE   HLA-A2/A28 family peptide A2.2F SEQ ID NO:107.
XX
KW   Major histocompatibility complex; MHC class I; MHC class II; antigen;
KW   immune response; diagnosis; antibody; immunisation; autoimmune disease;
KW   acquired immune deficiency syndrome; AIDS; cytostatic; dermatological;
KW   anti-inflammatory; neuroprotective; immunosuppressive; antithyroid;
KW   vaccine; lupus erythematosus; multiple sclerosis; thyroiditis;
KW   toxic shock; tumour; snakebite.
XX
OS   Mammalia.
XX
PN   US5976551-A.
XX
PD   02-NOV-1999.
XX
PF   07-JUN-1995;   95US-0484905.
XX
PR   05-DEC-1991;   91US-0801818.
PR   15-NOV-1991;   91US-0792473.
XX
PA   (INSP ) INST PASTEUR.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Kourilsky P,  Mottez E,  Abastado J;
XX
DR   WPI; 2000-037081/03.
XX
PT   Composition containing an antigen and altered major histocompatibility
PT   Class II determinant, used to immunize against autoimmune diseases,
PT   e.g. acquired immune deficiency syndrome  -
XX
PS   Disclosure; Column 149-152; 96pp; English.
XX
CC   The present invention describes a composition capable of eliciting
CC   anti-major histocompatibility (MHC) antibodies. The composition
CC   comprises an antigen associated with an altered MHC Class II determinant
CC   (I) comprising alpha1, alpha2, beta1 and beta2 polypeptide domains
CC   encoded by a mammalian MHC Class II locus covalently linked to form a
CC   polypeptide (I) containing beta2, alpha2, alpha1 and beta1 domains in
CC   sequence. The resulting Antigen-MHC complex is recognizable by the T cell
CC   receptor. The compositions are used for immunisation against, or
CC   treatment of, a wide range of autoimmune diseases, e.g. acquired immune
CC   deficiency syndrome (AIDS), lupus erythematosus, multiple sclerosis,
CC   thyroiditis, toxic shock, tumour and snakebite, depending on the nature
CC   of antigen. (I) is also used to analyse functional interactions between
CC   the various domains and for targeting lymphocyte receptors. Antibodies
CC   against (I) are produced by usual methods of immunisation or cell fusion,
CC   and may be humanised by standard methods. These antibodies are useful for
CC   diagnosis (detection or purification of MHC gene products), therapy
CC   (neutralising MHC on cell surfaces) and in the study of MHC and cellular
CC   processes. AAZ33240 to AAZ33242 and AAY52840 to AAY52970 represent
CC   sequences used in the exemplification of the present invention.
XX
SQ   Sequence   274 AA;

  Query Match             33.2%;  Score 505;  DB 21;  Length 274;
  Best Local Similarity   39.9%;  Pred. No. 2.5e-39;
  Matches  110;  Conservative   43;  Mismatches  115;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |||:||| ||:      : |
Db       2 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASRRMEPRAPWIEQE-GPEYW 60

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
             :: :|      ||  |:   :| |:  ||||| : ||::  |    .|| :| |||
Db      61 DGETRKVKAHSQTHRVDLSTLRGYYNQSEAGSHTLQRMYGCDVGSDWRFLRGYHQYAYDG 120

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db     121 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQWRAYLEGTCVEWLRRYLENGKETL 179

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     180 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 238

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 275
           ||:| |  : || |:|||||| |:| || :|| : |
Db     239 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLPW 274


RESULT 15
AAB58690
ID   AAB58690 standard; protein; 274 AA.
XX
AC   AAB58690;
XX
DT   13-MAR-2001  (first entry)
XX
DE   HLA-A2/A28 protein #11.
XX
KW   Major histocompatibility complex; MHC class I; immune; snake bite;
KW   T cell mediated autoimmune disease; AIDS; lupus erythematosus;
KW   toxic shock.
XX
OS   Unidentified.
XX
PN   US6153408-A.
XX
PD   28-NOV-2000.
XX
PF   09-JAN-1995;   95US-0370476.
XX
PR   15-NOV-1991;   91US-0792473.
PR   07-SEP-1993;   93US-0117575.
```

PR 05-DEC-1991; 91US-0801818.
PR 07-JUN-1993; 93US-0072787.
XX
PA (INSP) INST PASTEUR.
PA (INRM) INST NAT SANTE & RECH MEDICAL.
XX
PI Abastado J, Kourilsky P, Casrouge A, Ojcius D, Lone Y, Mottez E;
XX
DR WPI; 2001-060089/07.
XX
PT New altered major histocompatibility complex (MHC) class I determinant
PT useful for eliciting an immune response and/or for immunizing against
PT or treating diseases, for example, multiple sclerosis, AIDS, toxic
PT shock or snake bite -
XX
PS Disclosure; Column 29-33; 105pp; English.
XX
CC The present invention relates to a major histocompatibility complex
CC (MHC) class I determinant, which has alpha_1 alpha_2 alpha_3 and
CC beta2-microglobulin polypeptide domains encoded by a mammalian MHC
CC class I locus. The MHC class I determinants are useful for activating
CC the immune system and presenting antigens to the immune system to
CC elicit an antigenic response. The MHC class I determinants are also
CC useful for treating diseases, e.g. T cell mediated autoimmune disease,
CC AIDS, lupus erythematosus, toxic shock or snake bite. The altered MHC
CC class I determinants and compositions containing antigens bound to
CC the determinants are useful in diagnostic applications, e.g. altered
CC determinants may be used to target lymphocyte receptors and the
CC resulting bound determinant can be assayed.
XX
SQ Sequence 274 AA;

Query Match 33.2%; Score 505; DB 22; Length 274;
Best Local Similarity 39.9%; Pred. No. 2.5e-39;
Matches 110; Conservative 43; Mismatches 115; Indels 8; Gaps 7;

```
Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |||:||| ||:      : |
Db       2 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASRRMEPRAPWIEQE-GPEYW 60

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
              :: :|       ||  |:   :| |:  ||||| : ||::  |     || :| |||
Db      61 DGETRKVKAHSQTHRVDLSTLRGYYNQSEAGSHTLQRMYGCDVGSDWRFLRGYHQYAYDG 120

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db     121 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQWRAYLEGTCVEWLRRYLENGKETL 179

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :   |   :||| |:    |||| ||::||  ||: | :| :   .:: |  :  | ||
Db     180 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 238

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIW 275
           ||:| |  : || |:|||||| |:| || :|| : |
Db     239 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLPW 274
```

Search completed: March 31, 2003, 14:08:45
Job time : 76 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 31, 2003, 14:07:04 ; Search time 28 Seconds
(without alignments)
290.026 Million cell updates/sec

Title: US-10-092-404-1
Perfect score: 1522
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1522	100.0	276	4	US-09-094-964-1	Sequence 1, Appli
2	1522	100.0	348	3	US-08-652-265-2	Sequence 2, Appli
3	1522	100.0	348	4	US-08-834-497A-2	Sequence 2, Appli
4	1522	100.0	348	4	US-09-503-444A-2	Sequence 2, Appli
5	1522	100.0	348	4	US-09-277-457-2	Sequence 2, Appli
6	1513	99.4	276	4	US-09-094-964-2	Sequence 2, Appli
7	1513	99.4	348	3	US-08-652-265-6	Sequence 6, Appli
8	1513	99.4	348	4	US-08-834-497A-6	Sequence 6, Appli
9	1513	99.4	348	4	US-09-503-444A-6	Sequence 6, Appli
10	1511	99.3	348	3	US-08-652-265-4	Sequence 4, Appli
11	1511	99.3	348	4	US-08-834-497A-4	Sequence 4, Appli
12	1511	99.3	348	4	US-09-503-444A-4	Sequence 4, Appli
13	1502	98.7	276	4	US-09-094-964-3	Sequence 3, Appli
14	1502	98.7	348	3	US-08-652-265-8	Sequence 8, Appli
15	1502	98.7	348	4	US-08-834-497A-8	Sequence 8, Appli
16	1502	98.7	348	4	US-09-503-444A-8	Sequence 8, Appli
17	522	34.3	361	3	US-08-652-265-22	Sequence 22, Appl
18	522	34.3	361	4	US-08-834-497A-22	Sequence 22, Appl
19	522	34.3	361	4	US-09-503-444A-22	Sequence 22, Appl
20	513	33.7	365	3	US-08-652-265-23	Sequence 23, Appl
21	513	33.7	365	4	US-08-834-497A-23	Sequence 23, Appl
22	513	33.7	365	4	US-09-503-444A-23	Sequence 23, Appl
23	505	33.2	274	2	US-08-484-905-107	Sequence 107, App
24	505	33.2	274	3	US-08-481-985B-107	Sequence 107, App
25	505	33.2	274	4	US-08-370-476-107	Sequence 107, App
26	505	33.2	341	3	US-08-890-719-38	Sequence 38, Appl
27	504	33.1	365	2	US-08-484-905-97	Sequence 97, Appl
28	504	33.1	365	3	US-08-481-985B-97	Sequence 97, Appl
29	504	33.1	365	4	US-08-370-476-97	Sequence 97, Appl
30	503	33.0	274	2	US-08-484-905-108	Sequence 108, App
31	503	33.0	274	3	US-08-481-985B-108	Sequence 108, App
32	503	33.0	274	4	US-08-370-476-108	Sequence 108, App
33	503	33.0	365	2	US-08-484-905-100	Sequence 100, App
34	503	33.0	365	3	US-08-481-985B-100	Sequence 100, App
35	503	33.0	365	4	US-08-370-476-100	Sequence 100, App
36	502	33.0	274	1	US-08-222-851-1	Sequence 1, Appli
37	502	33.0	365	2	US-08-484-905-99	Sequence 99, Appl
38	502	33.0	365	3	US-08-481-985B-99	Sequence 99, Appl
39	502	33.0	365	4	US-08-370-476-99	Sequence 99, Appl
40	501	32.9	274	2	US-08-484-905-106	Sequence 106, App
41	501	32.9	274	3	US-08-481-985B-106	Sequence 106, App
42	501	32.9	274	4	US-08-370-476-106	Sequence 106, App
43	501	32.9	365	2	US-08-484-905-98	Sequence 98, Appl
44	501	32.9	365	3	US-08-481-985B-98	Sequence 98, Appl
45	501	32.9	365	4	US-08-370-476-98	Sequence 98, Appl

ALIGNMENTS

RESULT 1
US-09-094-964-1
; Sequence 1, Application US/09094964
; Patent No. 6391852
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Bjorkman, Pamela J.
; APPLICANT: Schatzman, Randall C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
; TITLE OF INVENTION: AND IRON DEFICIENCY DISEASES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,964
; FILING DATE: June 12, 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/876,010
; FILING DATE: June 13, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0074-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 276 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-094-964-1

Query Match 100.0%; Score 1522; DB 4; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.4e-142;

```
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||||||||||||||||||||||||||||||||||||
Db     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 2
US-08-652-265-2
; Sequence 2, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-2

  Query Match           100.0%;  Score 1522;  DB 3;  Length 348;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||||||||||||||||||||||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 3
US-08-834-497A-2
; Sequence 2, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
```

```
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-834-497A-2

  Query Match             100.0%;  Score 1522;  DB 4;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||||||||||||||||||||||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 4
US-09-503-444A-2
; Sequence 2, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-2

  Query Match             100.0%;  Score 1522;  DB 4;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||||||||||||||||||||||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 5
US-09-277-457-2
; Sequence 2, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 10653/002001
;  CURRENT APPLICATION NUMBER: US/09/277,457
;  CURRENT FILING DATE:  1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-277-457-2

  Query Match             100.0%;  Score 1522;  DB 4;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-142;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202
```

```
Qy    181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy    241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
          |||||||||||||||||||||||||||||||||||
Db    263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 6
US-09-094-964-2
; Sequence 2, Application US/09094964
; Patent No. 6391852
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Bjorkman, Pamela J.
;    APPLICANT:  Schatzman, Randall C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;    TITLE OF INVENTION:  AND IRON DEFICIENCY DISEASES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/094,964
;      FILING DATE:  June 12, 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/876,010
;      FILING DATE:  June 13, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0074-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-094-964-2

  Query Match             99.4%;  Score 1513;  DB 4;  Length 276;
  Best Local Similarity   99.6%;  Pred. No. 1.1e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 60

Qy     61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy    121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
```

```
Qy    181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy    241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
          |||||||||||||||||||||||||||||||||||
Db    241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 7
US-08-652-265-6
; Sequence 6, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-6

  Query Match             99.4%;  Score 1513;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy     61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy    121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy    181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262
```

```
Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||||||||||||||||||||||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 8
US-08-834-497A-6
; Sequence 6, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-6

  Query Match             99.4%;  Score 1513;  DB 4;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||||||||||||||||||||||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 9
US-09-503-444A-6
; Sequence 6, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141                                                    ;
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-6

  Query Match             99.4%;  Score 1513;  DB 4;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-141;
```

```
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||||||||||||||||||||||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 10
US-08-652-265-4
; Sequence 4, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-4

  Query Match           99.3%;  Score 1511;  DB 3;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 2.3e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||||||||||||||||||||||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 11
US-08-834-497A-4
; Sequence 4, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
```

```
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-834-497A-4

  Query Match             99.3%;  Score 1511;  DB 4;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 2.3e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||||||||||||||||| ||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 12
US-09-503-444A-4
; Sequence 4, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-503-444A-4

  Query Match             99.3%;  Score 1511;  DB 4;  Length 348;
  Best Local Similarity   99.6%;  Pred. No. 2.3e-141;
  Matches  275;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||||||||||||||||| ||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 13
US-09-094-964-3
; Sequence 3, Application US/09094964
; Patent No. 6391852
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Bjorkman, Pamela J.
;    APPLICANT:  Schatzman, Randall C.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF IRON OVERLOAD DISEASES
;    TITLE OF INVENTION:  AND IRON DEFICIENCY DISEASES
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds, LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/094,964
;      FILING DATE:  June 12, 1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/876,010
;      FILING DATE:  June 13, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0074-999
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-094-964-3

  Query Match             98.7%;  Score 1502;  DB 4;  Length 276;
  Best Local Similarity   99.3%;  Pred. No. 1.3e-140;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNASKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 QDALEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||||||||||||||||||||||||||||||||||||
Db     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276


RESULT 14
US-08-652-265-8
; Sequence 8, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-652-265-8

  Query Match             98.7%;  Score 1502;  DB 3;  Length 348;
  Best Local Similarity   99.3%;  Pred. No. 1.8e-140;
  Matches  274;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||||||||||||||||| ||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298


RESULT 15
US-08-834-497A-8
; Sequence 8, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
```

; APPLICATION NUMBER: US 08/630,912
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-834-497A-8

Query Match 98.7%; Score 1502; DB 4; Length 348;
Best Local Similarity 99.3%; Pred. No. 1.8e-140;
Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDDESRRVEPRTPWVSSRISSQ 82

Qy     61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy    121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy    181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy    241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
          |||||||||||||||||| ||||||||||||||||||
Db    263 TYQGWITLAVPPGEEQRYTYQVEHPGLDQPLIVIWE 298
```

Search completed: March 31, 2003, 14:10:25
Job time : 30 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 31, 2003, 14:09:13 ; Search time 35 Seconds
(without alignments)
463.078 Million cell updates/sec

Title: US-10-092-404-1
Perfect score: 1522
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	514	33.8	92	9	US-10-016-634A-120	Sequence 120, App
2	505	33.2	280	9	US-10-073-300-6	Sequence 6, Appli
3	505	33.2	415	9	US-10-073-300-5	Sequence 5, Appli
4	476	31.3	542	9	US-10-015-535-32	Sequence 32, Appl
5	476	31.3	542	9	US-10-015-535-34	Sequence 34, Appl
6	475	31.2	542	9	US-10-015-535-36	Sequence 36, Appl
7	473	31.1	540	9	US-10-015-535-22	Sequence 22, Appl
8	473	31.1	541	9	US-10-015-535-28	Sequence 28, Appl
9	473	31.1	542	9	US-10-015-535-24	Sequence 24, Appl
10	473	31.1	542	9	US-10-015-535-26	Sequence 26, Appl
11	447	29.4	332	10	US-09-870-521-3	Sequence 3, Appli
12	444	29.2	540	9	US-10-015-535-30	Sequence 30, Appl
13	443	29.1	334	10	US-09-870-521-4	Sequence 4, Appli
14	358.5	23.6	170	10	US-09-925-301-1307	Sequence 1307, Ap
15	335	22.0	271	10	US-09-925-301-1431	Sequence 1431, Ap
16	275	18.1	145	10	US-09-810-560-8	Sequence 8, Appli
17	242	15.9	184	9	US-09-858-580-21	Sequence 21, Appl
18	242	15.9	184	9	US-09-847-172-21	Sequence 21, Appl
19	226	14.8	91	10	US-09-864-761-38005	Sequence 38005, A
20	223	14.7	91	10	US-09-864-761-35461	Sequence 35461, A
21	210.5	13.8	104	10	US-09-925-302-835	Sequence 835, App
22	207	13.6	117	10	US-09-810-560-9	Sequence 9, Appli
23	196.5	12.9	93	10	US-09-864-761-39479	Sequence 39479, A
24	196.5	12.9	110	10	US-09-864-761-35339	Sequence 35339, A
25	196.5	12.9	114	10	US-09-864-761-37988	Sequence 37988, A
26	174.5	11.5	261	9	US-09-925-664-30	Sequence 30, Appl
27	174	11.4	411	9	US-10-015-536-17	Sequence 17, Appl
28	173	11.4	110	9	US-09-796-692-799	Sequence 799, App
29	173	11.4	110	9	US-09-796-692-2139	Sequence 2139, Ap
30	171.5	11.3	285	10	US-09-756-983-24	Sequence 24, Appl
31	167	11.0	772	10	US-09-815-837-74	Sequence 74, Appl
32	166.5	10.9	448	12	US-10-081-281-111	Sequence 111, App
33	166	10.9	246	9	US-09-992-598-225	Sequence 225, App
34	166	10.9	246	9	US-09-989-293A-225	Sequence 225, App
35	166	10.9	246	9	US-09-989-735-225	Sequence 225, App
36	166	10.9	246	9	US-09-990-444-225	Sequence 225, App
37	166	10.9	246	9	US-09-989-730-225	Sequence 225, App
38	166	10.9	246	9	US-09-990-436-225	Sequence 225, App
39	166	10.9	246	9	US-09-991-181-225	Sequence 225, App
40	166	10.9	246	9	US-09-993-687-225	Sequence 225, App
41	166	10.9	246	9	US-09-989-734-225	Sequence 225, App
42	166	10.9	246	9	US-10-028-072-436	Sequence 436, App
43	166	10.9	246	9	US-09-997-653-225	Sequence 225, App
44	166	10.9	246	9	US-10-174-590-600	Sequence 600, App
45	166	10.9	246	9	US-10-176-758-600	Sequence 600, App

ALIGNMENTS

RESULT 1
US-10-016-634A-120
; Sequence 120, Application US/10016634A
; Publication No. US20020192666A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Ghosh, Malavika
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and
; FILE REFERENCE: DEX-0255
; CURRENT APPLICATION NUMBER: US/10/016,634A
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: US 60/244,258
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 120
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-016-634A-120

Query Match 33.8%; Score 514; DB 9; Length 92;
Best Local Similarity 100.0%; Pred. No. 4.5e-40;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      92 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHK 151
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHK 60

Qy     152 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
           |||||||||||||||||||||||||||||||
Db      61 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 92
```

RESULT 2
US-10-073-300-6
; Sequence 6, Application US/10073300
; Publication No. US20030003535A1
; GENERAL INFORMATION:
; APPLICANT: Reiter, Yoram

```
;  TITLE OF INVENTION: SINGLE CHAIN CLASS I MAJOR HISTO- COMPATIBILITY COMPLEXES
;  FILE REFERENCE: 02/23339
;  CURRENT APPLICATION NUMBER: US/10/073,300
;  CURRENT FILING DATE:  2002-06-25
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 280
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-073-300-6

  Query Match             33.2%;  Score 505;  DB 9;  Length 280;
  Best Local Similarity   39.4%;  Pred. No. 1.3e-38;
  Matches  109;  Conservative   45;  Mismatches  115;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db       2 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 60

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
           :: :|      ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db      61 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 120

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||   : | | ||||  |   |:: || |:  |
Db     121 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGTCVEWLRRYLENGKETL 179

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     180 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 238

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |:|||||| |:| || :|| : ||
Db     239 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 275


RESULT 3
US-10-073-300-5
; Sequence 5, Application US/10073300
; Publication No. US20030003535A1
; GENERAL INFORMATION:
;  APPLICANT: Reiter, Yoram
;  TITLE OF INVENTION: SINGLE CHAIN CLASS I MAJOR HISTO- COMPATIBILITY COMPLEXES
;  FILE REFERENCE: 02/23339
;  CURRENT APPLICATION NUMBER: US/10/073,300
;  CURRENT FILING DATE:  2002-06-25
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 415
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: human beta2 microglobulin linked to MHC class I heavy chain
US-10-073-300-5

  Query Match             33.2%;  Score 505;  DB 9;  Length 415;
  Best Local Similarity   39.4%;  Pred. No. 2.1e-38;
  Matches  109;  Conservative   45;  Mismatches  115;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db     117 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 175

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
           :: :|      ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db     176 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 235

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||   : | | ||||  |   |:: || |:  |
Db     236 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGTCVEWLRRYLENGKETL 294

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     295 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 353

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |:|||||| |:| || :|| : ||
Db     354 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 390


RESULT 4
US-10-015-535-32
; Sequence 32, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-32

  Query Match             31.3%;  Score 476;  DB 9;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 1.3e-35;
  Matches  109;  Conservative   41;  Mismatches  118;  Indels    8;  Gaps    7;

Qy       6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
           ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db      56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
           : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db     115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy     122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
           |::    |   | ||:  |  || :||:    | : |||||  |  :|:: |: |   |
Db     175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVERLRRYLKNGNATLL 233

Qy     182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db     234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |:| | :: || |:|| ||| | | || :|| : ||
Db     293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 5
US-10-015-535-34
; Sequence 34, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
```

```
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-34

  Query Match             31.3%;  Score 476;  DB 9;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 1.3e-35;
  Matches  109;  Conservative   41;  Mismatches  118;  Indels    8;  Gaps    7;

Qy       6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
           ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db      56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
           : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db     115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy     122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
           |::    |   | ||:  |  || :||:    | : |||||  |  :|:: |: |   |
Db     175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVERLRRYLKNGNATLL 233

Qy     182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db     234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |:| | :: || |:|| ||| | | || :|| : ||
Db     293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 6
US-10-015-535-36
; Sequence 36, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-36

  Query Match             31.2%;  Score 475;  DB 9;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 1.6e-35;
  Matches  109;  Conservative   41;  Mismatches  118;  Indels    8;  Gaps    7;

Qy       6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFY--DHESRRVEPRTPWVSSRISSQMWL 63
           ||| |      |   ||   :  :|||||  || :  | |: | |||  |:      : |
Db      56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFNSDAENPRYEPRARWMEQE-GPEYWE 114

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
           : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db     115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy     122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
           |::    |   | ||:  |  || :||:    | : |||||  |  :|:: |: |   |
Db     175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVERLRRYLKNGNATLL 233

Qy     182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db     234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |:| | :: || |:|| ||| | | || :|| : ||
Db     293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 7
US-10-015-535-22
; Sequence 22, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 540
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-22

  Query Match             31.1%;  Score 473;  DB 9;  Length 540;
  Best Local Similarity   39.5%;  Pred. No. 2.5e-35;
  Matches  109;  Conservative   40;  Mismatches  119;  Indels    8;  Gaps    7;

Qy       6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
           ||| |      |   ||   :  :|||||  || :|   |: | |||  |:      : |
Db     144 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 202

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
           : :|  || :  | ||  |::  :| ||  |||:||| |||:  |     || :| |||
Db     203 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 262

Qy     122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
           |::    |   | ||:  |  || :||:    | : |||||  |  :|:: |: |   |
Db     263 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVEWLRRYLKNGNATLL 321

Qy     182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db     322 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 380

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |:| | :: || |:|| ||| | | || :|| : ||
Db     381 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 416


RESULT 8
US-10-015-535-28
; Sequence 28, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
```

```
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 541
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-28

  Query Match             31.1%;  Score 473;  DB 9;  Length 541;
  Best Local Similarity   39.5%;  Pred. No. 2.5e-35;
  Matches  109;  Conservative   40;  Mismatches  119;  Indels    8;  Gaps    7;

Qy       6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
             ||| |      |   ||   :  :|||||  || :|   |: | |||  |:       : |
Db      56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
           : :|  || :  | ||  |::  :| ||  |||:||| |||:  |      || :| |||
Db     115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy     122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
           |::    |   | ||:  |  || :||:    | : |||||  |   |:: |: |   |
Db     175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVEWLRRYLKNGNATLL 233

Qy     182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db     234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |:| | :: || |:|| ||| | | || :|| : ||
Db     293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 9
US-10-015-535-24
; Sequence 24, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-24

  Query Match             31.1%;  Score 473;  DB 9;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 2.5e-35;
  Matches  109;  Conservative   40;  Mismatches  119;  Indels    8;  Gaps    7;

Qy       6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
             ||| |      |   ||   :  :|||||  || :|   |: | |||  |:       : |
Db      56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
           : :|  || :  | ||  |::  :| ||  |||:||| |||:  |      || :| |||
Db     115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy     122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
           |::    |   | ||:  |  || :||:    | : |||||  |   |:: |: |   |
Db     175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVEWLRRYLKNGNATLL 233

Qy     182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db     234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |:| | :: || |:|| ||| | | || :|| : ||
Db     293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 10
US-10-015-535-26
; Sequence 26, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-26

  Query Match             31.1%;  Score 473;  DB 9;  Length 542;
  Best Local Similarity   39.5%;  Pred. No. 2.5e-35;
  Matches  109;  Conservative   40;  Mismatches  119;  Indels    8;  Gaps    7;

Qy       6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
             ||| |      |   ||   :  :|||||  || :|   |: | |||  |:       : |
Db      56 HSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQE-GPEYWE 114

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
           : :|  || :  | ||  |::  :| ||  |||:||| |||:  |      || :| |||
Db     115 RETQKAKGNEQSFRVDLRTLLGYYNQSKGGSHTIQVISGCEVGSDGRLLRGYQQYAYDGC 174

Qy     122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
           |::    |   | ||:  |  || :||:    | : |||||  |   |:: |: |   |
Db     175 DYIALNEDLKTWTAADMAALITKHKWEQAG-EAERLRAYLEGTCVEWLRRYLKNGNATLL 233

Qy     182 QQVPPLVKVTHHV-TSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
           :   |   ||||       |||| || :|| :||: |  : : :  :: |  :  | |||
Db     234 RTDSPKAHVTHHSRPEDKVTLRCWALGFYPADITLTWQLNGEEL-IQDMELVETRPAGDG 292

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |:| | :: || |:|| ||| | | || :|| : ||
Db     293 TFQKWASVVVPLGKEQYYTCHVYHQGLPEPLTLRWE 328


RESULT 11
US-09-870-521-3
; Sequence 3, Application US/09870521
; Patent No. US20020051989A1
; GENERAL INFORMATION:
;  APPLICANT: Miller, Marcia
```

```
;  APPLICANT:  Goto, Ronald
;  TITLE OF INVENTION: METHOD FOR BREEDING AND GENOTYPING CHICKENS AND PROBES THEREFOR
;  FILE REFERENCE: 1954-310
;  CURRENT APPLICATION NUMBER: US/09/870,521
;  CURRENT FILING DATE:  2001-06-01
;  PRIOR APPLICATION NUMBER: US 60/208471
;  PRIOR FILING DATE: 2000-06-02
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 332
;   TYPE: PRT
;   ORGANISM: Gallus sp.
US-09-870-521-3

  Query Match             29.4%;  Score 447;  DB 10;  Length 332;
  Best Local Similarity   35.3%;  Pred. No. 3.2e-33;
  Matches   96;  Conservative   49;  Mismatches  125;  Indels    2;  Gaps    2;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQ 64
Db       2 SHSLRYFLTGMTDPGPGMPRFVIVGYVDDKIFGTYNSKSRTAQPIVEMLPQE-DQEHWDT 60

Qy      65 LSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHL 124
Db      61 QTQKAQGGERDFDWNLNRLPERYNKSKGSHTMQMMFGCDILEDGSIRGYDQYAFDGRDFL 120

Qy     125 EFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQV 184
Db     121 AFDMDTMTFTAADPVAEITKRRWETEGTYAERWKHELGTVCVQNLRRYLEHGKAALKRRV 180

Qy     185 PPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQG 244
Db     181 QPEVRVWGKEADGILTLSCHAHGFYPRPITISWMKDGMVRD-QETRWGGIVPNSDGTYHA 239

Qy     245 WITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
Db     240 SAAIDVLPEDGDKYWCRVEHASLPQPGLFSWE 271


RESULT 12
US-10-015-535-30
; Sequence 30, Application US/10015535
; Publication No. US20030036506A1
; GENERAL INFORMATION:
;  APPLICANT: Kranz, David M.
;  APPLICANT:  Brophy, Susan
;  TITLE OF INVENTION: Mutated Class I Major Histocompatibility proteins and
;  TITLE OF INVENTION:  Complexes
;  FILE REFERENCE: 100-00
;  CURRENT APPLICATION NUMBER: US/10/015,535
;  CURRENT FILING DATE:  2001-12-10
;  PRIOR APPLICATION NUMBER: 60/254,495
;  PRIOR FILING DATE: 2000-12-08
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
;   LENGTH: 540
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: peptide
US-10-015-535-30

  Query Match             29.2%;  Score 444;  DB 9;  Length 540;
  Best Local Similarity   35.9%;  Pred. No. 1.1e-32;
  Matches   99;  Conservative   52;  Mismatches  117;  Indels    8;  Gaps    7;

Qy       6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDH--ESRRVEPRTPWVSSRISSQMWL 63
Db     144 HSMRYFETAVSRRGLGEPRYISVGYVNDKEFVRFDSDAENPRYEPRAPWMEQE-GPEYWE 202

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSK-ESHTLQVILGCEMQEDNS-TEGYWKYGYDGQ 121
Db     203 RITQIAKGQEQWFRVNLRTLLGYYNQSAGGTHTLQWMYGCDVGSDGRLLRGYEQFAYDGC 262

Qy     122 DHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLD 181
Db     263 DYIALNEDLKTWTFADMSSMITRRKWEQAG-AAEYYRAYLEGECVEWLHRYLKNGNATLL 321

Qy     182 QQVPPLVKVTHHVTS-SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
Db     322 RTDSPKAHVTYHPRSKGEVTLRCWALGFYPADITLTWQLNGEEL-TQDMELVETRPAGDG 380

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
Db     381 TFQKWASVVVPLGKEQNYTCRVYHEGLPHPLTLRWE 416


RESULT 13
US-09-870-521-4
; Sequence 4, Application US/09870521
; Patent No. US20020051989A1
; GENERAL INFORMATION:
;  APPLICANT: Miller, Marcia
;  APPLICANT:  Goto, Ronald
;  TITLE OF INVENTION: METHOD FOR BREEDING AND GENOTYPING CHICKENS AND PROBES THEREFO
;  FILE REFERENCE: 1954-310
;  CURRENT APPLICATION NUMBER: US/09/870,521
;  CURRENT FILING DATE:  2001-06-01
;  PRIOR APPLICATION NUMBER: US 60/208471
;  PRIOR FILING DATE: 2000-06-02
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 334
;   TYPE: PRT
;   ORGANISM: Gallus sp.
US-09-870-521-4

  Query Match             29.1%;  Score 443;  DB 10;  Length 334;
  Best Local Similarity   34.2%;  Pred. No. 7.5e-33;
  Matches   93;  Conservative   53;  Mismatches  122;  Indels    4;  Gaps    4;

Qy       6 HSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQL 65
Db       3 HTLRYIQTAMTDPGPGQPWFVTVGYVDGELFVHYNSTARRYVPRTEWIAAKAQEQ-YDTG 61

Qy      66 SQSLKGWDHMFTVDFWTIMENHN-HSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHL 124
Db      62 TQKIGGGNRQIDRELNGIPQRYNKQTGGSHTVQWMYGCDILEGGPIRGYYQMAYDGRDFT 121

Qy     125 EFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQV 184
Db     122 AFDKGTMTFTAAVPEAVPTKRKWEEESEPERW-KNYLEETCVEWLRRYVEYGKAELGRRE 180

Qy     185 PPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQG 244
Db     181 RPEVRVWGKEADGILTLSCRAHGFYPRPIVVSWLKD-GAVRGQDAHSGGIVPNGDGTYHT 239

Qy     245 WITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
Db     240 WVTIDAQPGDGDKYQCRVEHASLPQPGLYSWE 271


RESULT 14
US-09-925-301-1307
; Sequence 1307, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
```

```
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1307
;   LENGTH: 170
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-301-1307

  Query Match             23.6%;  Score 358.5;  DB 10;  Length 170;
  Best Local Similarity   40.5%;  Pred. No. 1.7e-25;
  Matches   68;  Conservative   32;  Mismatches   59;  Indels    9;  Gaps    3;

Qy    113 YWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQL 172
          :||| |||:|::|| :   |   :| |  || :||   :  :: :|||| :||| |::
Db      6 FWKYYYDGKDYIEFNKEIPAWVPFDPAAQITKQKWEAEPVYVQRAKAYLEEECPATLRKY 65

Qy    173 LELGRGVLDQQVPPLVKVT-HHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEP 231
          |:  : :||:| || | || |        |:| | ::||  | : | :  :       ||
Db     66 LKYSKNILDRQDPPSVVVTSHQAPGEKKKLKCLAYDFYPGKIDVHWTRAGEVQ-----EP 120

Qy    232 K---DVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
          :   ||| ||:|||| |: :|||| :   |:| |:|  | |||:| ||
Db    121 ELRGDVLHNGNGTYQSWVVVAVPPQDTAPYSCHVQHSSLAQPLVVPWE 168


RESULT 15
US-09-925-301-1431
; Sequence 1431, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1431
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-301-1431

  Query Match             22.0%;  Score 335;  DB 10;  Length 271;
  Best Local Similarity   34.6%;  Pred. No. 4.5e-23;
  Matches   85;  Conservative   39;  Mismatches  110;  Indels   12;  Gaps    8;

Qy      5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESR--RVEPRTPWVSSRISSQMW 62
          |||| |     |    |   : |: |||| |: :| ::   |:||| |||      | |
Db     29 SHSLRYFSTAVSRPGRGEPRYIAVEYVDDTQFLRFDSDAAIPRMEPREPWVEQE-GPQYW 87

Qy     63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
             :  |       |    ::  :| |:  ||||| : ||:|  |     || :: |||
Db     88 EWTTGYAKANAQTDRVALRNLLRRYNQSEAGSHTLQGMNGCDMGPDGRLLRGYHQHAYDG 147

Qy    121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
          :|::    |   | ||:  |  |:  :|  :  | : | ||| :|   |:: || |:  |
Db    148 KDYISLNEDLRSWTAADTVAQITQRFYEAEEY-AEEFRTYLEGECLELLRRYLENGKETL 206

Qy    181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKD--KQPMDAKEFEPKDVLPN 237
          :  ||   | || ::    |||| || :||  ||: | :|  :|  | :  | :   |
Db    207 QRADPPKAHVAHHPISDHEATLRCWALGFYPAEITLTWQRDGEEQTQDTELVETR---PA 263

Qy    238 GDGTYQ 243
          ||||::
Db    264 GDGTFR 269


Search completed: March 31, 2003, 14:14:15
Job time : 38 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 31, 2003, 14:07:03 ; Search time 11 Seconds
(without alignments)
1040.679 Million cell updates/sec

Title: US-10-092-404-1
Perfect score: 1522
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1522	100.0	348	1	HFE_HUMAN	Q30201 homo sapien
2	1165	76.5	360	1	HFE_RAT	O35799 rattus norv
3	1149	75.5	359	1	HFE_MOUSE	P70387 mus musculu
4	522	34.3	361	1	HA1A_RABIT	P01894 oryctolagus
5	522	34.3	361	1	HA1B_RABIT	P06140 oryctolagus
6	516	33.9	365	1	1A01_PANTR	P16209 pan troglod
7	515	33.8	364	1	HA1B_BOVIN	P13753 bos taurus
8	513	33.7	365	1	1A11_HUMAN	P13746 homo sapien
9	511	33.6	370	1	1A03_HUMAN	P04439 homo sapien
10	509	33.4	365	1	1A80_HUMAN	Q09160 homo sapien
11	507	33.3	365	1	1A31_HUMAN	P16189 homo sapien
12	505	33.2	365	1	1A02_HUMAN	P01892 homo sapien
13	505	33.2	365	1	1A30_HUMAN	P16188 homo sapien
14	505	33.2	365	1	1A74_HUMAN	P30459 homo sapien
15	503	33.0	365	1	1A03_PANTR	P13748 pan troglod
16	502	33.0	365	1	1A33_HUMAN	P16190 homo sapien
17	502	33.0	365	1	1A36_HUMAN	P30455 homo sapien
18	502	33.0	365	1	1A68_HUMAN	P01891 homo sapien
19	500.5	32.9	362	1	HA19_CANFA	P18466 canis famil
20	500	32.9	365	1	1A01_HUMAN	P30443 homo sapien
21	499	32.8	273	1	1A69_HUMAN	P10316 homo sapien
22	499	32.8	365	1	1A04_PANTR	P13749 pan troglod
23	499	32.8	365	1	1A24_HUMAN	P05534 homo sapien
24	497	32.7	360	1	HA1A_BOVIN	P13752 bos taurus
25	496	32.6	296	1	ZA2G_RAT	Q63678 rattus norv
26	496	32.6	362	1	1B45_HUMAN	P30485 homo sapien
27	495	32.5	365	1	1A23_HUMAN	P30447 homo sapien
28	493	32.4	338	1	1B20_HUMAN	P30467 homo sapien
29	492	32.3	363	1	1B04_GORGO	P30382 gorilla gor
30	491	32.3	295	1	ZA2G_HUMAN	P25311 homo sapien
31	491	32.3	322	1	HA10_MOUSE	P01898 mus musculu
32	491	32.3	362	1	1B29_HUMAN	P18463 homo sapien
33	491	32.3	371	1	HA12_RAT	P16391 rattus norv
34	490	32.2	365	1	1A34_HUMAN	P30453 homo sapien
35	490	32.2	365	1	1A66_HUMAN	P30457 homo sapien
36	489	32.1	338	1	HLAG_HUMAN	P17693 homo sapien
37	489	32.1	361	1	1B14_HUMAN	P03989 homo sapien
38	489	32.1	362	1	1B18_HUMAN	P10318 homo sapien
39	489	32.1	366	1	1C02_GORGO	P30385 gorilla gor
40	489	32.1	366	1	1C04_GORGO	P30387 gorilla gor
41	488	32.1	359	1	1B01_PANTR	P13750 pan troglod
42	488	32.1	362	1	1B16_HUMAN	P19373 homo sapien
43	488	32.1	362	1	1B19_HUMAN	Q08136 homo sapien
44	488	32.1	365	1	1A01_PONPY	P16211 pongo pygma
45	488	32.1	365	1	1A32_HUMAN.	P10314 homo sapien

ALIGNMENTS

RESULT 1
HFE_HUMAN
ID HFE_HUMAN STANDARD; PRT; 348 AA.
AC Q30201; O75929; O75930; O75931;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hereditary hemochromatosis protein precursor (HLA-H).
GN HFE OR HLAH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS HH ASP-63 AND TYR-282.
RX MEDLINE=96331279; PubMed=8696333;
RA Feder J.N., Gnirke A., Thomas W., Tsuchihashi Z., Ruddy D.A.,
RA Basava A., Dormishian F., Domingo R., Ellis M.C. Jr., Fullan A.,
RA Hinton L.M., Jones N.L., Kimmel B.E., Kronmal G.S., Lauer P.,
RA Lee V.K., Loeb D.B., Mapa F.A., McClelland E., Meyer N.C.,
RA Mintier G.A., Moeller N., Moore T., Morikang E., Prass C.E.,
RA Quintana L., Starnes S.M., Schatzman R.C., Brunke K.J.,
RA Drayna D.T., Risch N.J., Bacon B.R., Wolff R.K.;
RT "A novel MHC class I-like gene is mutated in patients with hereditary
RT haemochromatosis.";
RL Nat. Genet. 13:399-409(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Albig W., Burmester N., Bode C., Doenecke D., Drabent B.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97294057; PubMed=9149941;
RA Ruddy D.A., Kronmal G.S., Lee V.K., Mintier G.A., Quintana L.,
RA Domingo R. Jr., Meyer N.C., Irrinki A., McClelland E.E., Fullan A.,
RA Mapa F.A., Moore T., Thomas W., Loeb D.B., Harmon C., Tsuchihashi Z.,
RA Wolff R.K., Schatzman R.C., Feder J.N.;
RT "A 1.1-Mb transcript map of the hereditary hemochromatosis locus.";
RL Genome Res. 7:441-456(1997).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Gasparini P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3 AND 4).
RX MEDLINE=99180629; PubMed=10079302;
RA Rhodes D.A., Trowsdale J.;
RT "Alternate splice variants of the hemochromatosis gene Hfe.";
RL Immunogenetics 49:357-359(1999).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Oliva R., Sanchez M.;
RT "Identification of different alternative splicing forms of the HFE
RT gene.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN [7]

```
RP   FUNCTION.
RX   MEDLINE=98132614; PubMed=9465039;
RA   Feder J.N., Penny D.M., Irrinki A., Lee V.K., Lebron J.A., Watson N.,
RA   Tsuchihashi Z., Sigal E., Bjorkman P.J., Schatzman R.C.;
RT   "The hemochromatosis gene product complexes with the transferrin
RT   receptor and lowers its affinity for ligand binding.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:1472-1477(1998).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX   MEDLINE=98206473; PubMed=9546397;
RA   Lebron J.A., Bennett M.J., Vaughn D.E., Chirino A.J., Snow P.M.,
RA   Mintier G.A., Feder J.N., Bjorkman P.J.;
RT   "Crystal structure of the hemochromatosis protein HFE and
RT   characterization of its interaction with transferrin receptor.";
RL   Cell 93:111-123(1998).
RN   [9]
RP   VARIANTS HH ASP-63 AND TYR-282.
RX   MEDLINE=97260408; PubMed=9106528;
RA   Carella M., D'Ambrosio L., Totaro A., Grifa A., Valentino M.A.,
RA   Piperno A., Girelli D., Roetto A., Franco B., Gasparini P.,
RA   Camaschella C.;
RT   "Mutation analysis of the HLA-H gene in Italian hemochromatosis
RT   patients.";
RL   Am. J. Hum. Genet. 60:828-832(1997).
RN   [10]
RP   VARIANT HH/PCT TYR-282.
RX   MEDLINE=97176837; PubMed=9024376;
RA   Roberts A.G., Whatley S.D., Morgan R.R., Worwood M., Elder G.H.;
RT   "Increased frequency of the haemochromatosis Cys282Tyr mutation in
RT   sporadic porphyria cutanea tarda.";
RL   Lancet 349:321-323(1997).
RN   [11]
RP   VARIANT HH/PCT ASP-63.
RX   MEDLINE=98085904; PubMed=9425935;
RA   Sampietro M., Piperno A., Lupica L., Arosio C., Vergani A.,
RA   Corbetta N., Malosio I., Mattioli M., Fracanzani A.L.,
RA   Cappellini M.D., Fiorelli G., Fargion S.;
RT   "High prevalence of the His63Asp HFE mutation in Italian patients with
RT   porphyria cutanea tarda.";
RL   Hepatology 27:181-184(1998).
RN   [12]
RP   VARIANTS HH/PCT ASP-63 AND TYR-282.
RX   MEDLINE=98281650; PubMed=9620340;
RA   Bonkovsky H.L., Poh-Fitzpatrick M., Pimstone N., Obando J.,
RA   Di Bisceglie A., Tattrie C., Tortorelli K., LeClair P., Mercurio M.G.,
RA   Lambrecht R.W.;
RT   "Porphyria cutanea tarda, hepatitis C, and HFE gene mutations in North
RT   America.";
RL   Hepatology 27:1661-1669(1998).
RN   [13]
RP   VARIANTS HH ASP-63; CYS-65 AND TYR-282.
RX   MEDLINE=99211934; PubMed=10194428;
RA   Mura C., Raguenes O., Ferec C.;
RT   "HFE mutations analysis in 711 hemochromatosis probands: evidence for
RT   S65C implication in mild form of hemochromatosis.";
RL   Blood 93:2502-2505(1999).
RN   [14]
RP   VARIANTS HH CYS-65; ARG-93 AND THR-105.
RX   MEDLINE=20042794; PubMed=10575540;
RA   Barton J.C., Sawada-Hirai R., Rothenberg B.E., Acton R.T.;
RT   "Two novel missense mutations of the HFE gene (I105T and G93R) and
RT   identification of the S65C mutation in Alabama hemochromatosis
RT   probands.";
RL   Blood Cells Mol. Dis. 25:147-155(1999).
RN   [15]
RP   VARIANTS VP D-63 AND H-127, VARIANT HH M-330, AND VARIANTS M-53 AND
RP   M-59.
RX   MEDLINE=99330560; PubMed=10401000;
RA   de Villiers J.N.P., Hillermann R., Loubser L., Kotze M.J.;
RT   "Spectrum of mutations in the HFE gene implicated in haemochromatosis
RT   and porphyria.";
RL   Hum. Mol. Genet. 8:1517-1522(1999).
RN   [16]
RP   VARIANTS HH ASP-63 AND TYR-282.
RX   MEDLINE=99140260; PubMed=10094552;
RA   Merryweather-Clarke A.T., Simonsen H., Shearman J.D., Pointon J.J.,
RA   Norgaard-Pedersen B., Robson K.J.H.;
RT   "A retrospective anonymous pilot study in screening newborns for HFE
RT   mutations in Scandinavian populations.";
RL   Hum. Mutat. 13:154-159(1999).
RN   [17]
RP   VARIANT HH CYS-65.
RA   Fagan E., Payne S.J.;
RT   "A novel missense mutation S65C in the HFE gene with a possible role
RT   in hereditary haemochromatosis.";
RL   Hum. Mutat. 13:507-508(1999).
RN   [18]
RP   VARIANT LYS-277.
RX   MEDLINE=20081073; PubMed=10612845;
RA   Bradbury R., Fagan E., Payne S.J.;
RT   "Two novel polymorphisms (E277K and V212V) in the haemochromatosis
RT   gene HFE.";
RL   Hum. Mutat. 15:120-120(2000).
CC   -!- FUNCTION: BINDS TO TRANSFERRIN RECEPTOR (TFR) AND REDUCES ITS
CC       AFFINITY FOR IRON-LOADED TRANSFERRIN.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2/delE2,
CC       3/del14E4 and 4/delE214E4; are produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: IN ALL TISSUES TESTED EXCEPT BRAIN.
CC   -!- DISEASE: DEFECTS IN HFE ARE A CAUSE OF HEREDITARY HEMOCHROMATOSIS
CC       (HH). HH IS AN AUTOSOMAL RECESSIVE INBORN DISORDER OF IRON
CC       METABOLISM, FREQUENT AMONG CAUCASIANS. HH IS CHARACTERIZED BY
CC       ABNORMAL INTESTINAL IRON ABSORPTION AND PROGRESSIVE INCREASE OF
CC       TOTAL BODY IRON, WHICH RESULTS IN MIDLIFE IN CLINICAL
CC       COMPLICATIONS INCLUDING CIRRHOSIS, CARDIOPATHY, DIABETES,
CC       ENDOCRINE DYSFUNCTIONS, ARTHROPATHY, AND SUSCEPTIBILITY TO LIVER
CC       CANCER. SINCE THE DISEASE COMPLICATIONS CAN BE EFFECTIVELY
CC       PREVENTED BY REGULAR PHLEBOTOMIES, EARLY DIAGNOSIS IS MOST
CC       IMPORTANT TO PROVIDE A NORMAL LIFE EXPECTANCY TO THE AFFECTED
CC       SUBJECTS.
CC   -!- DISEASE: DEFECTS IN HFE ARE A CAUSE OF PORPHYRIA CUTANEA
CC       TARDA (PCT), A DISORDER CHARACTERIZED BY LIGHT-SENSITIVE
CC       DERMATITIS AND PRESENCE OF LARGE AMOUNTS OF UROPORPHYRIN IN
CC       URINE. IRON OVERLOAD IS OFTEN PRESENT IN ASSOCIATION WITH VARYING
CC       DEGREES OF LIVER DAMAGE. PCT IS THE MOST COMMON FORM OF PORPHYRIA
CC       WORLDWIDE. IT OCCURS IN TWO FORMS: THE SPORADIC TYPE (PCT TYPE I)
CC       AND THE FAMILIAL TYPE (PCT TYPE II), WHICH IS INHERITED IN AN
CC       AUTOSOMAL DOMINANT MANNER.
CC   -!- DISEASE: DEFECTS IN HFE ARE A CAUSE OF VARIEGATE PORPHYRIA (VP),
CC       THE MOST COMMON FORM OF PORPHYRIA IN SOUTH AFRICA. THIS AUTOSOMAL
CC       DOMINANT DISEASE IS CHARACTERIZED BY SKIN HYPERPIGMENTATION AND
CC       HYPERTRICHOSIS, ABDOMINAL PAIN, TACHYCARDIA, HYPERTENSION AND
CC       NEUROMUSCULAR DISTURBANCES. HIGH FECAL LEVELS OF PROTOPORPHYRIN
CC       AND COPROPORPHYRIN, INCREASED URINE UROPORPHYRINS AND IRON
CC       OVERLOAD ARE TYPICAL MARKERS OF THE DISEASE.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U60319; AAC51823.1; -.
DR   EMBL; Z92910; CAB07442.1; -.
DR   EMBL; U91328; AAB82083.1; -.
DR   EMBL; Y09801; CAA70934.1; -.
DR   EMBL; Y09800; CAA70934.1; JOINED.
DR   EMBL; Y09803; CAA70934.1; JOINED.
DR   EMBL; Y09799; CAA70934.1; JOINED.
DR   EMBL; AF079407; AAC62646.1; -.
DR   EMBL; AF079408; AAC62647.1; -.
DR   EMBL; AF079409; AAC62648.1; -.
DR   EMBL; AJ249336; CAC67793.1; -.
```

```
  Query Match             100.0%;  Score 1522;  DB 1;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-123;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 202

Qy     181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 262

Qy     241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
          ||||||||||||||||||||||||||||||||||||
Db     263 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 298


RESULT 2
HFE_RAT
ID   HFE_RAT        STANDARD;      PRT;   360 AA.
AC   O35799; O35175;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein homolog precursor (RT1-CAFE).
GN   HFE.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Banasch M.W., Schaefer H., Schmidt W.E.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 42-303 FROM N.A.
RC   TISSUE=Small intestine;
RA   Sawada-Hirai R., Rothenberg B.E.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: BINDS TO TRANSFERRIN RECEPTOR (TFR) AND REDUCES ITS
CC       AFFINITY FOR IRON-LOADED TRANSFERRIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ001517; CAA04799.1; -.
DR   EMBL; AF008587; AAB86597.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    360       HEREDITARY HEMOCHROMATOSIS PROTEIN
FT                                HOMOLOG.
FT   DOMAIN       26    127       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      128    218       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      219    310       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      311    319       CONNECTING PEPTIDE.
FT   TRANSMEM    320    340       POTENTIAL.
FT   DOMAIN      341    360       CYTOPLASMIC TAIL.
FT   DISULFID    137    200       BY SIMILARITY.
FT   DISULFID    238    295       BY SIMILARITY.
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    143    143       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    167    167       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    247    247       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    198    198       R -> K (IN REF. 2).
SQ   SEQUENCE   360 AA;  40988 MW;  CC819834EE3240B3 CRC64;

  Query Match             76.5%;  Score 1165;  DB 1;  Length 360;
  Best Local Similarity   73.9%;  Pred. No. 2.3e-92;
  Matches  207;  Conservative   29;  Mismatches   36;  Indels    8;  Gaps    1;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQ 64
          |||| |||||||: ||||  |||||||||||||| |:||||| ||| ||:  : |||:|||
Db      32 SHSLRYLFMGASKPDLGLPFFEALGYVDDQLFVSYNHESRRAEPRAPWILGQTSSQLWLQ 91

Qy      65 LSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWKY 116
          |||||||||:|| ||||||| |:||||        ||| |||||||: ||||| |:|||
Db      92 LSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLRVVPESHILQVILGCEVHEDNSTSGFWKY 151

Qy     117 GYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELG 176
          |||||||||||| ||:| |||||| ||:||| |:||||:| ||:|||| ||:|:|||
Db     152 GYDGQDHLEFCPKTLNWSAAEPRAWATKMEWEEHRIRARQSRDYLQRDCPQQLKQVLELQ 211

Qy     177 RGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
          |||| |||| ||||| |  |: |:|||:|||::|||||:|||| ||:|||:  |::|||
Db     212 RGVLGQQVPTLVKVTRHWASTGTSLRCQALNFFPQNITMRWLKDSQPLDAKDVNPENVLP 271

Qy     237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
          |||||||||:|||| |||| |::||||||||||||   ||
Db     272 NGDGTYQGWLTLAVAPGEETRFSCQVEHPGLDQPLTATWE 311


RESULT 3
HFE_MOUSE
ID   HFE_MOUSE      STANDARD;      PRT;   359 AA.
AC   P70387;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Hereditary hemochromatosis protein homolog precursor.
GN   HFE OR MR2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=98060831; PubMed=9396865;
RA   Riegert P., Gilfillan S., Nanda I., Schmid M., Bahram S.;
RT   "The mouse HFE gene.";
RL   Immunogenetics 47:174-177(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Lung;
RA   Hashimoto K.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 37-211 FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=97148566; PubMed=9020055;
```

```
RA   Hashimoto K., Hirai M., Kurosawa Y.;
RT   "Identification of a mouse homolog for the human hereditary
RT   haemochromatosis candidate gene.";
RL   Biochem. Biophys. Res. Commun. 230:35-39(1997).
RN   [4]
RP   SEQUENCE OF 79-359 FROM N.A.
RC   STRAIN=129;
RA   Albig W., Drabent B., Doenecke D.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: BINDS TO TRANSFERRIN RECEPTOR (TFR) AND REDUCES ITS
CC       AFFINITY FOR IRON-LOADED TRANSFERRIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: TO MHC CLASS I ANTIGENS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF007558; AAC03447.1; -.
DR   EMBL; U66849; AAB07525.1; -.
DR   EMBL; Y12650; CAA73197.1; -.
DR   EMBL; U80604; AAB51504.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   MGD; MGI:109191; Hfe.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    359       HEREDITARY HEMOCHROMATOSIS PROTEIN
FT                                HOMOLOG.
FT   DOMAIN       25    126       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      127    217       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      218    309       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      310    318       CONNECTING PEPTIDE.
FT   TRANSMEM    319    339       POTENTIAL.
FT   DOMAIN      340    359       CYTOPLASMIC TAIL.
FT   DISULFID    136    199       BY SIMILARITY.
FT   DISULFID    237    294       BY SIMILARITY.
FT   CARBOHYD    114    114       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    142    142       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   359 AA;  40548 MW;  4BDE6C27F9FF20B4 CRC64;

  Query Match             75.5%;  Score 1149;  DB 1;  Length 359;
  Best Local Similarity   72.6%;  Pred. No. 5.3e-91;
  Matches  204;  Conservative   30;  Mismatches   39;  Indels    8;  Gaps    1;

Qy       4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
           ||||| |||||||| |||| |||| ||||||||| |:||||| ||| ||:  : |||:||
Db      30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWK 115
           |||||||||:|| ||||||| |:|||         ||| |||:|||| ||||| |:|:
Db      90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy     116 YGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
           |||||||||||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db     150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy     176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
           ||||| |||| ||||| |  |: |:|||:||:::|||||:|||| ||:|||:  |: ||
Db     210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy     236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||| ||||||:|||| ||:| |:|||||||||||||   ||
Db     270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310


RESULT 4
HA1A_RABIT
ID   HA1A_RABIT     STANDARD;      PRT;   361 AA.
AC   P01894;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-JAN-1990 (Rel. 13, Last annotation update)
DE   RLA class I histocompatibility antigen, alpha chain 11/11 precursor.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84290724; PubMed=6432910;
RA   Tykocinski M.L., Marche P.N., Max E.E., Kindt T.J.;
RT   "Rabbit class I MHC genes: cDNA clones define full-length transcripts
RT   of an expressed gene and a putative pseudogene.";
RL   J. Immunol. 133:2261-2269(1984).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02441; AAA98729.1; -.
DR   PIR; A02193; HLRB.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    361       RLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                ALPHA CHAIN 11/11.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    329
FT   DOMAIN      330    361       CYTOPLASMIC.
FT   SIMILAR     207    298       SIGNIFICANT WITH IMMUNOGLOBULIN C-REGION
FT                                DOMAINS AND BETA-2-MICROGLOBULIN.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
SQ   SEQUENCE   361 AA;  40447 MW;  580B673323C1AE35 CRC64;

  Query Match             34.3%;  Score 522;  DB 1;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 2e-37;
  Matches  111;  Conservative   44;  Mismatches  114;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | :   |   ||   |  :|||||  || :| :  | |:| | ||:  ::  : |
Db      26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84
```

```
Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | :|  |      | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db      85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           |::    |   | ||:  |  |: :||     | ::||||||:|   |:: ||:|:  |
Db     145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy     181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   ||||   |    |||| || :||  |:: | :| :    :: |  :  | ||
Db     204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || ||||||||:|:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299


RESULT 5
HA1B_RABIT
ID   HA1B_RABIT     STANDARD;      PRT;   361 AA.
AC   P06140;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   01-JAN-1990 (Rel. 13, Last annotation update)
DE   RLA class I histocompatibility antigen, alpha chain 19-1 precursor.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85103547; PubMed=3917974;
RA   Marche P.N., Tykocinski M.L., Max E.E., Kindt T.J.;
RT   "Structure of a functional rabbit class I MHC gene: similarity to
RT   human class I genes.";
RL   Immunogenetics 21:71-82(1985).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02819; AAA98730.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL       1     24
FT   CHAIN       25    361       RLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 ALPHA CHAIN 19-1.
FT   DOMAIN      25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN     115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN     207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN     299    308       CONNECTING PEPTIDE.
FT   TRANSMEM   309    329
FT   DOMAIN     330    361       CYTOPLASMIC.
FT   SIMILAR    207    298       SIGNIFICANT WITH IMMUNOGLOBULIN C-REGION
FT                                 DOMAINS AND BETA-2-MICROGLOBULIN.
FT   CARBOHYD   110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID   125    188       BY SIMILARITY.
FT   DISULFID   227    283       BY SIMILARITY.
SQ   SEQUENCE   361 AA;  40455 MW;  C06FBD8B87ED0546 CRC64;

  Query Match             34.3%;  Score 522;  DB 1;  Length 361;
  Best Local Similarity   40.1%;  Pred. No. 2e-37;
  Matches  111;  Conservative   44;  Mismatches  114;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | :   |   ||   |  :|||||  || :| :  | |:| | ||:  ::  : |
Db      26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | :|  |      | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db      85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           |::    |   | ||:  |  |: :||     | ::||||||:|   |:: ||:|:  |
Db     145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy     181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   ||||   |    |||| || :||  |:: | :| :    :: |  :  | ||
Db     204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || ||||||||:|:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299


RESULT 6
1A01_PANTR
ID   1A01_PANTR     STANDARD;      PRT;   365 AA.
AC   P16209;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   CHLA class I histocompatibility antigen, A-2 alpha chain precursor.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90201944; PubMed=1690682;
RA   Lawlor D.A., Warren E., Ward F.E., Parham P.;
RT   "Comparison of class I MHC alleles in humans and apes.";
RL   Immunol. Rev. 113:147-185(1990).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30678; AAA87970.1; -.
DR   HSSP; Q95352; 1HHK.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL       1     24
FT   CHAIN       25    365       CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
```

```
FT                                A-2 ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365       CYTOPLASMIC TAIL.
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   365 AA;  40848 MW;  FC452786BD038D3E CRC64;

  Query Match             33.9%;  Score 516;  DB 1;  Length 365;
  Best Local Similarity   39.7%;  Pred. No. 6.6e-37;
  Matches  110;  Conservative   45;  Mismatches  114;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           : ::| |       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db      85 DEETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQRRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |||||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 7
HA1B_BOVIN
ID   HA1B_BOVIN     STANDARD;      PRT;   364 AA.
AC   P13753;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-JAN-1990 (Rel. 13, Last annotation update)
DE   BOLA class I histocompatibility antigen, alpha chain BL3-7 precursor.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88258075; PubMed=3133413;
RA   Ennis P.D., Jackson A.P., Parham P.;
RT   "Molecular cloning of bovine class I MHC cDNA.";
RL   J. Immunol. 141:642-651(1988).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21043; AAA30641.1; -.
DR   PIR; B27638; B27638.
DR   HSSP; P16391; 1ED3.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     27
FT   CHAIN        28    364       BOLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 ALPHA CHAIN BL3-7.
FT   DOMAIN       28    117       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      118    209       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      210    301       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      302    310       CONNECTING PEPTIDE.
FT   TRANSMEM    311    331
FT   DOMAIN      332    364       CYTOPLASMIC.
FT   SIMILAR     210    301       SIGNIFICANT WITH IMMUNOGLOBULIN C-REGION
FT                                 DOMAINS AND BETA-2-MICROGLOBULIN.
FT   CARBOHYD    106    106       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    128    191       BY SIMILARITY.
FT   DISULFID    230    286       BY SIMILARITY.
SQ   SEQUENCE   364 AA;  41513 MW;  622056CF7DCF7873 CRC64;

  Query Match             33.8%;  Score 515;  DB 1;  Length 364;
  Best Local Similarity   38.9%;  Pred. No. 8e-37;
  Matches  109;  Conservative   50;  Mismatches  113;  Indels    8;  Gaps    7;

Qy       2 LLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISS 59
           |  |||| | : | |   ||   | |:|||||  || :| :  : | ||| ||:
Db      26 LAGSHSLRYFYTGVSRPGLGEPRFIAVGYVDDTQFVRFDSDAPNPREEPRVPWMEQE-GP 84

Qy      60 QMWLQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYG 117
           : | : ::  |    :| ||  |:   :| |:  || :| : ||::  |     |:|::|
Db      85 EYWDRNTRIYKDTAQIFRVDLNTLRGYYNQSETGSHNIQAMYGCDVGPDGRLLRGFWQFG 144

Qy     118 YDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGR 177
           |||:|::    :   | ||:  |  || :||     |   | ||| :|   |:: || |:
Db     145 YDGRDYIALNEELRSWTAADTAAQITKRKWEAAG-AAETWRNYLEGECVEWLRRYLENGK 203

Qy     178 GVLDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
           | :  ||   |||| ::    |||| || :||: |:: | :: :    :: |  :  |
Db     204 DTLLRADPPKAHVTHHSISDREVTLRCWALGFYPEEISLTWQREGED-QTQDMELVETRP 262

Qy     237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           :||||:| |  | || ||||||||:|:| || :|| : ||
Db     263 SGDGTFQKWAALVVPSGEEQRYTCRVQHEGLQEPLTLRWE 302


RESULT 8
1A11_HUMAN
ID   1A11_HUMAN     STANDARD;      PRT;   365 AA.
AC   P13746;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   HLA class I histocompatibility antigen, A-11 alpha chain precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*1101/A*1102).
RX   MEDLINE=89030641; PubMed=2460344;
RA   Mayer W.E., Jonker M., Klein D., Ivanyi P., van Seventer G.,
RA   Klein J.;
RT   "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for
RT   trans-species mode of evolution.";
RL   EMBO J. 7:2765-2774(1988).
RN   [2]
RP   SEQUENCE FROM N.A. (A*1101/A*1102).
```

```
RX   MEDLINE=94287401; PubMed=8016845;
RA   Lin L., Tokunaga K., Ishikawa Y., Bannai M., Kashiwase K.,
RA   Kuwata S., Akaza T., Tadokoro K., Shibata Y., Juji T.;
RT   "Sequence analysis of serological HLA-A11 split antigens, A11.1 and
RT   A11.2.";
RL   Tissue Antigens 43:78-82(1994).
RN   [3]
RP   SEQUENCE OF 26-365 FROM N.A. (A*1101).
RX   MEDLINE=87192928; PubMed=2437024;
RA   Cowan E.P., Jelachich M.L., Biddison W.E., Coligan J.E.;
RT   "DNA sequence of HLA-A11: remarkable homology with HLA-A3 allows
RT   identification of residues involved in epitopes recognized by
RT   antibodies and T cells.";
RL   Immunogenetics 25:241-250(1987).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-11 ARE KNOWN: A*1101 (A-
CC       11E) AND A*1102 (A-11K). THE SEQUENCE SHOWN IS THAT OF A*1101.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X13111; CAA31503.1; -.
DR   EMBL; X13112; CAA31504.1; -.
DR   EMBL; D16841; BAA04117.1; -.
DR   EMBL; D16842; BAA04118.1; -.
DR   EMBL; M16010; AAA65449.1; -.
DR   EMBL; M16007; AAA65449.1; JOINED.
DR   EMBL; M16008; AAA65449.1; JOINED.
DR   EMBL; M16009; AAA65449.1; JOINED.
DR   PIR; S03536; S03536.
DR   PIR; S03694; S03694.
DR   PIR; A47636; A47636.
DR   HSSP; O19673; 1HSB.
DR   MIM; 142800; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL       1     24
FT   CHAIN       25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-11 ALPHA CHAIN.
FT   DOMAIN      25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN     115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN     207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN     299    308       CONNECTING PEPTIDE.
FT   TRANSMEM   309    332
FT   DOMAIN     333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD   110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID   125    188       BY SIMILARITY.
FT   DISULFID   227    283       BY SIMILARITY.
FT   VARIANT     43     43       E -> K (IN A*1102).
FT                               /FTId=VAR_004353.
SQ   SEQUENCE   365 AA;  40937 MW;  FE449CE2D4BF6CC5 CRC64;

  Query Match             33.7%;  Score 513;  DB 1;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 1.2e-36;
  Matches  109;  Conservative   47;  Mismatches  113;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | :   |    |   | |:|||||  || :| :  |:|:||| ||:      : |

Db      26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
             | ::::|       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db      85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |   : || ||||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 9
1A03_HUMAN
ID   1A03_HUMAN     STANDARD;      PRT;   370 AA.
AC   P04439;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   HLA class I histocompatibility antigen, A-3 alpha chain precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*0301).
RX   MEDLINE=84207948; PubMed=6609814;
RA   Strachan T., Sodoyer R., Damotte M., Jordan B.R.;
RT   "Complete nucleotide sequence of a functional class I HLA gene,
RT   HLA-A3: implications for the evolution of HLA genes.";
RL   EMBO J. 3:887-894(1984).
RN   [2]
RP   SEQUENCE FROM N.A. (A*0302).
RX   MEDLINE=85290871; PubMed=2993417;
RA   Cowan E.P., Jordan B.E., Coligan J.E.;
RT   "Molecular cloning and DNA sequence analysis of genes encoding
RT   cytotoxic T lymphocyte-defined HLA-A3 subtypes: the E1 subtype.";
RL   J. Immunol. 135:2835-2841(1985).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-3 ARE KNOWN: A*0301 (A-
CC       3.1) AND A*0302. THE SEQUENCE SHOWN IS THAT OF A*0301.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X00492; CAA25162.1; ALT_TERM.
DR   PIR; A02192; HLHUA3.
DR   HSSP; O19673; 1HSB.
DR   MIM; 142800; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
```

```
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     29
FT   CHAIN        30    370       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-3 ALPHA CHAIN.
FT   DOMAIN       30    119       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      120    211       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      212    303       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      304    313       CONNECTING PEPTIDE.
FT   TRANSMEM    314    337
FT   DOMAIN      338    370       CYTOPLASMIC TAIL.
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    130    193       BY SIMILARITY.
FT   DISULFID    232    288       BY SIMILARITY.
FT   VARIANT     181    181       E -> V (IN A*0302).
FT                                /FTId=VAR_004351.
FT   VARIANT     185    185       L -> Q (IN A*0302).
FT                                /FTId=VAR_004352.
SQ   SEQUENCE   370 AA;  41368 MW;  ABB1FA77460318A2 CRC64;

  Query Match             33.6%;  Score 511;  DB 1;  Length 370;
  Best Local Similarity   39.6%;  Pred. No. 1.8e-36;
  Matches  110;  Conservative   47;  Mismatches  111;  Indels   10;  Gaps    8;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      31 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 89

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
            | ::::|       ||  |:   :| |:  |||:|:: ||::  |     || :  |||
Db      90 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 149

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWE-RHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
           :|::    |   | ||:  |  || :||  |:  | | ||||:  |   |:: || |:
Db     150 KDYIALNEDLRSWTAADMAAQITKRKWEAAHE--AEQLRAYLDGTCVEWLRRYLENGKET 207

Qy     180 LDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNG 238
           | :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | |
Db     208 LQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAG 266

Qy     239 DGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||:| |  : || |||||||| |:| || :|| : ||
Db     267 DGTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 304


RESULT 10
1A80_HUMAN
ID   1A80_HUMAN     STANDARD;      PRT;   365 AA.
AC   Q09160;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   HLA class I histocompatibility antigen, AW-80(A-1) alpha chain
DE   precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94245293; PubMed=8188325;
RA   Balas A., Garcia-Sanchez F., Gomez-Reino F., Vicario J.L.;
RT   "Characterization of a new and highly distinguishable HLA-A allele in
RT   a Spanish family.";
RL   Immunogenetics 39:452-452(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Domena J.D.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC     THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   -!- POLYMORPHISM: THE ONLY ALLELE OF AW-80 KNOWN IS A*8001 WHICH IS
CC       SHOWN HERE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U03754; AAC04322.1; -.
DR   EMBL; L18898; AAA17012.1; -.
DR   HSSP; Q95352; 1HHK.
DR   MIM; 142800; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 AW-80(A-1) ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
SQ   SEQUENCE   365 AA;  40791 MW;  CE1BC1CD60CA8FA8 CRC64;

  Query Match             33.4%;  Score 509;  DB 1;  Length 365;
  Best Local Similarity   38.3%;  Pred. No. 2.6e-36;
  Matches  106;  Conservative   53;  Mismatches  110;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDSQFVQFDSDAASQRMEPRAPWIEQE-EPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
            : ::::|       :  |:   :| |:: |||:|:: ||::  |     || :  |||
Db      85 DEETRNVKAHSQTNRANLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||  : || | |||||  :|   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITKRKWEAAR-RAEQLRAYLEGECVDGLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   :||| ::    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |:|:|||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGKEKRYTCHVQHEGLPEPLTLRWE 299


RESULT 11
1A31_HUMAN
ID   1A31_HUMAN     STANDARD;      PRT;   365 AA.
AC   P16189; O98137; Q9TQ24;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   HLA class I histocompatibility antigen, A-31 alpha chain precursor
DE   (Aw-19).
GN   HLA-A OR HLAA.
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*3101).
RX   MEDLINE=90038496; PubMed=2478623;
RA   Kato K., Trapani J.A., Allopenna J., Dupont B., Yang S.Y.;
RT   "Molecular analysis of the serologically defined HLA-Aw19 antigens. A
RT   genetically distinct family of HLA-A antigens comprising A29, A31,
RT   A32, and Aw33, but probably not A30.";
RL   J. Immunol. 143:3371-3378(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (A*3101).
RX   MEDLINE=92269955; PubMed=1317015;
RA   Belich M.P., Madrigal J.A., Hildebrand W.H., Zemmour J.,
RA   Williams R.C., Luz R., Petzl-Erler M.L., Parham P.;
RT   "Unusual HLA-B alleles in two tribes of Brazilian Indians.";
RL   Nature 357:326-329(1992).
RN   [3]
RP   SEQUENCE FROM N.A. (A*31012).
RX   MEDLINE=96387675; PubMed=8795145;
RA   Arnett K.L., Adams E.J., Parham P.;
RT   "On the sequence of A*3101.";
RL   Tissue Antigens 47:428-430(1996).
RN   [4]
RP   SEQUENCE OF 9-365 FROM N.A. (A*3101).
RX   MEDLINE=92269956; PubMed=1589035;
RA   Watkins D.I., McAdam S.N., Liu X., Stang C.R., Milford E.L.,
RA   Levine C.G., Garber T.L., Dogon A.L., Lord C.I., Ghim S.H.,
RA   Troup G.M., Hughes A.L., Letvin N.L.;
RT   "New recombinant HLA-B alleles in a tribe of South American
RT   Amerindians indicate rapid evolution of MHC class I loci.";
RL   Nature 357:329-333(1992).
RN   [5]
RP   SEQUENCE FROM N.A. (A*3104).
RA   Bettinotti M.P., Dhillon G., Hackett J., Simonis T.B., Marincola F.M.;
RT   "A New HLA-A*31 allele.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 26-206 FROM N.A. (A*3104).
RA   Mitsuishi Y.;
RT   "New HLA-A31 allele identified in African American population.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-31 ARE KNOWN: A*3101 AND
CC       A*3104. THE SEQUENCE SHOWN IS THAT OF A*3101.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30578; AAA59613.1; -.
DR   EMBL; M84375; AAA59599.1; -.
DR   EMBL; L78918; AAB05976.1; -.
DR   EMBL; AF148863; AAD39981.1; -.
DR   EMBL; AF105028; AAC79721.1; -.
DR   EMBL; AF105027; AAC79721.1; JOINED.
DR   HSSP; O19673; 1HSB.
DR   MIM; 142800; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     24
FT   CHAIN        25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-31 ALPHA CHAIN.
FT   DOMAIN        25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN       115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN       207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN       299    308       CONNECTING PEPTIDE.
FT   TRANSMEM     309    332
FT   DOMAIN       333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD     110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID     125    188       BY SIMILARITY.
FT   DISULFID     227    283       BY SIMILARITY.
FT   VARIANT      121    121       M -> I (IN A*3104).
FT                                 /FTId=VAR_010373.
FT   VARIANT      138    138       Q -> R (IN A*3104).
FT                                 /FTId=VAR_010374.
SQ   SEQUENCE   365 AA;  41004 MW;  4E760C821A3C553B CRC64;

  Query Match             33.3%;  Score 507;  DB 1;  Length 365;
  Best Local Similarity   39.0%;  Pred. No. 3.9e-36;
  Matches  108;  Conservative   51;  Mismatches  110;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: |     |   |   | |:|||||  || :| :  |:|:||| ||:       : |
Db      26 SHSMRYFTTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-RPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
             | ::::|    :  ||  |:   :| |:  |||:|:: ||::  |     || :  |||
Db      85 DQETRNVKAHSQIDRVDLGTLRGYYNQSEAGSHTIQMMYGCDVGSDGRFLRGYQQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  |: :||  :: | | |||||  |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITQRKWEAARV-AEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :  ||   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| | :: || |:|||||| |:| || :|| : ||
Db     263 GTFQKWASVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 12
1A02_HUMAN
ID   1A02_HUMAN     STANDARD;      PRT;   365 AA.
AC   P01892; P06338; P30514; P30444; P30445; P30446; Q29680; Q29899;
AC   Q95352; Q29837; Q95380;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   HLA class I histocompatibility antigen, A-2 alpha chain precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*0201).
RX   MEDLINE=85132727; PubMed=2982951;
RA   Koller B.H., Orr H.T.;
RT   "Cloning and complete sequence of an HLA-A2 gene: analysis of two
RT   HLA-A alleles at the nucleotide level.";
RL   J. Immunol. 134:2727-2733(1985).
RN   [2]
RP   SEQUENCE FROM N.A. (A*0201).
RX   MEDLINE=89122144; PubMed=2914713;
RA   Cianetti L., Testa U., Scotto L., la Valle R., Simeone A.,
RA   Boccoli G., Giannella G., Peschle C., Boncinelli E.;
RT   "Three new class I HLA alleles: structure of mRNAs and alternative
```

RT mechanisms of processing.";
RL Immunogenetics 29:80-91(1989).
RN [3]
RP SEQUENCE FROM N.A. (A*0201).
RX MEDLINE=90207291; PubMed=2320591;
RA Ennis P.D., Zemmour J., Salter R.D., Parham P.;
RT "Rapid cloning of HLA-A,B cDNA by using the polymerase chain
RT reaction: frequency and nature of errors produced in amplification.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2833-2837(1990).
RN [4]
RP SEQUENCE FROM N.A. (A*0201/A*0211/A*0212).
RX MEDLINE=92269955; PubMed=1317015;
RA Belich M.P., Madrigal J.A., Hildebrand W.H., Zemmour J.,
RA Williams R.C., Luz R., Petzl-Erler M.L., Parham P.;
RT "Unusual HLA-B alleles in two tribes of Brazilian Indians.";
RL Nature 357:326-329(1992).
RN [5]
RP SEQUENCE OF 39-365 FROM N.A. (A*0201).
RX MEDLINE=85230571; PubMed=3874058;
RA Krangel M.S.;
RT "Unusual RNA splicing generates a secreted form of HLA-A2 in a
RT mutagenized B lymphoblastoid cell line.";
RL EMBO J. 4:1205-1210(1985).
RN [6]
RP SEQUENCE OF 25-295 (A*0201).
RX MEDLINE=80056745; PubMed=92029;
RA Orr H.T., Lopez de Castro J.A., Parham P., Ploegh H.L.,
RA Strominger J.L.;
RT "Comparison of amino acid sequences of two human histocompatibility
RT antigens, HLA-A2 and HLA-B7: location of putative alloantigenic
RT sites.";
RL Proc. Natl. Acad. Sci. U.S.A. 76:4395-4399(1979).
RN [7]
RP REVISIONS (A*0201).
RX MEDLINE=82247941; PubMed=6179086;
RA Lopez de Castro J.A., Strominger J.L., Strong D.M., Orr H.T.;
RT "Structure of crossreactive human histocompatibility antigens HLA-A28
RT and HLA-A2: possible implications for the generation of HLA
RT polymorphism.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:3813-3817(1982).
RN [8]
RP SEQUENCE OF 26-298 FROM N.A. (A*0202/A*0203).
RX MEDLINE=87306734; PubMed=3497874;
RA Mattson D.H., Handy D.E., Bradley D.A., Coligan J.E., Cowan E.P.,
RA Biddison W.E.;
RT "DNA sequences of the genes that encode the CTL-defined HLA-A2
RT variants M7 and DK1.";
RL Immunogenetics 26:190-192(1987).
RN [9]
RP SEQUENCE FROM N.A. (A*0203/A*0205).
RX MEDLINE=87252273; PubMed=3496393;
RA Holmes N., Ennis P., Wan A.M., Denney D.W., Parham P.;
RT "Multiple genetic mechanisms have contributed to the generation of
RT the HLA-A2/A28 family of class I MHC molecules.";
RL J. Immunol. 139:936-941(1987).
RN [10]
RP SEQUENCE FROM N.A. (A*0203/A*0205).
RA Domena J.D.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
RN [11]
RP SEQUENCE OF 9-365 FROM N.A. (A*0204).
RX MEDLINE=92039809; PubMed=1937577;
RA Castano A.R., Lopez de Castro J.A.;
RT "Structure of the HLA-A*0204 antigen, found in South American
RT Indians. Spatial clustering of HLA-A2 subtype polymorphism.";
RL Immunogenetics 34:281-285(1991).
RN [12]
RP SEQUENCE OF 9-365 FROM N.A. (A*0204).
RX MEDLINE=92269956; PubMed=1589035;
RA Watkins D.I., McAdam S.N., Liu X., Stang C.R., Milford E.L.,
RA Levine C.G., Garber T.L., Dogon A.L., Lord C.I., Ghim S.H.,
RA Troup G.M., Hughes A.L., Letvin N.L.;
RT "New recombinant HLA-B alleles in a tribe of South American
RT Amerindians indicate rapid evolution of MHC class I loci.";
RL Nature 357:329-333(1992).
RN [13]
RP SEQUENCE FROM N.A. (A*0206).
RX MEDLINE=89235215; PubMed=2715640;
RA Parham P., Lawlor D.A., Lomen C.E., Ennis P.D.;
RT "Diversity and diversification of HLA-A,B,C alleles.";
RL J. Immunol. 142:3937-3950(1989).
RN [14]
RP PARTIAL SEQUENCE (A*0206).
RX MEDLINE=86305811; PubMed=3489037;
RA Ezquerra A., Domenech N., van der Poel J., Strominger J.L., Vega M.A.,
RA Lopez de Castro J.A.;
RT "Molecular analysis of an HLA-A2 functional variant CLA defined by
RT cytolytic T lymphocytes.";
RL J. Immunol. 137:1642-1649(1986).
RN [15]
RP PARTIAL SEQUENCE (A*0207).
RX MEDLINE=88113844; PubMed=2448239;
RA Domenech N., Ezquerra A., Castano R., Lopez de Castro J.A.;
RT "Structural analysis of HLA-A2.4 functional variant KNE. Implications
RT for the mapping of HLA-A2-specific T-cell epitopes.";
RL Immunogenetics 27:196-202(1988).
RN [16]
RP PARTIAL SEQUENCE (A*0208).
RX MEDLINE=88314183; PubMed=2457548;
RA Domenech N., Castano R., Goulmy E., Lopez de Castro J.A.;
RT "Molecular analysis of HLA-A2.4 functional variant KLO: close
RT structural and evolutionary relatedness to the HLA-A2.2 subtype.";
RL Immunogenetics 28:143-152(1988).
RN [17]
RP PARTIAL SEQUENCE (A*0209).
RX MEDLINE=88186100; PubMed=3258580;
RA Castano R., Ezquerra A., Domenech N., Lopez de Castro J.A.;
RT "An HLA-A2 population variant with structural polymorphism in the
RT alpha 3 region.";
RL Immunogenetics 27:345-355(1988).
RN [18]
RP SEQUENCE FROM N.A. (A*0210).
RX MEDLINE=89122133; PubMed=2783680;
RA Epstein H., Kennedy L., Holmes N.;
RT "An Oriental HLA-A2 subtype is closely related to a subset of
RT Caucasoid HLA-A2 alleles.";
RL Immunogenetics 29:112-116(1989).
RN [19]
RP SEQUENCE OF 9-365 FROM N.A. (A*0211).
RX MEDLINE=92218010; PubMed=1559719;
RA Castano A.R., Lopez de Castro J.A.;
RT "Structure of the HLA-A*0211 (A2.5) subtype: further evidence for
RT selection-driven diversification of HLA-A2 antigens.";
RL Immunogenetics 35:344-346(1992).
RN [20]
RP SEQUENCE FROM N.A. (A*0213).
RX MEDLINE=94222455; PubMed=8168863;
RA Barber D.F., Fernandez J.M., Lopez de Castro J.A.;
RT "Primary structure of a new HLA-A2 subtype: HLA-A*0213.";
RL Immunogenetics 39:378-378(1994).
RN [21]
RP SEQUENCE FROM N.A. (A*0216).
RX MEDLINE=95278976; PubMed=7759139;
RA Barouch D., Krausa P., Bodmer J., Browning M.J., McMichael A.J.;
RT "Identification of a novel HLA-A2 subtype, HLA-A*0216.";
RL Immunogenetics 41:388-388(1995).
RN [22]
RP SEQUENCE FROM N.A. (A*0217).
RC TISSUE=Blood;
RX MEDLINE=95381236; PubMed=7652742;
RA Selvakumar A., Granja C.B., Salazar M., Alosco S.M., Yunis E.J.,
RA Dupont B.;
RT "A novel subtype of A2 (A*0217) isolated from the South American
RT Indian B-cell line AMALA.";
RL Tissue Antigens 45:343-347(1995).
RN [23]

```
RP   SEQUENCE FROM N.A. (A*0218).
RC   TISSUE=Blood;
RA   Kashiwase K., Tokunaga K., Ishikawa Y., Oohashi H., Hashimoto M.,
RA   Akaza T., Tadokoro K., Juji T.;
RT   "A new A2 sequence HLA-A2K from Japanese.";
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
RN   [24]
RP   SEQUENCE FROM N.A. (A*0220).
RC   TISSUE=Blood;
RX   MEDLINE=97161038; PubMed=9008310;
RA   Fleischhauer K., Zino E., Mazzi B., Severini G.M., Benazzi E.,
RA   Bordignon C.;
RT   "HLA-A*02 subtype distribution in Caucasians from northern Italy:
RT   identification of A*0220.";
RL   Tissue Antigens 48:673-679(1996).
RN   [25]
RP   SEQUENCE FROM N.A. (A*0221).
RC   TISSUE=Blood;
RA   Szmania S., Baxter-Lowe L.A.;
RT   "Nucleotide sequence of a novel HLA-A2 gene.";
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [26]
RP   X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF A*0201.
RX   MEDLINE=88014204; PubMed=3309677;
RA   Bjorkman P.J., Saper M.A., Samraoui B., Bennett W.S.,
RA   Strominger J.L., Wiley D.C.;
RT   "Structure of the human class I histocompatibility antigen, HLA-A2.";
RL   Nature 329:506-512(1987).
RN   [27]

  Query Match             33.2%;  Score 505;  DB 1;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 5.8e-36;
  Matches  109;  Conservative   45;  Mismatches  115;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
             :: :|       ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db      85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||   : | | ||||| |   |:: || |:  |
Db     145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :   |   :||| |:    |||| ||::|| ||: | :| :    :: |  :  | ||
Db     204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |:|||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 13
1A30_HUMAN
ID   1A30_HUMAN     STANDARD;      PRT;   365 AA.
AC   P16188; P30452; Q9UIP7;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   HLA class I histocompatibility antigen, A-30(AW-19) alpha chain
DE   precursor.
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*3001).
RX   MEDLINE=90038496; PubMed=2478623;
RA   Kato K., Trapani J.A., Allopenna J., Dupont B., Yang S.Y.;
RT   "Molecular analysis of the serologically defined HLA-Aw19 antigens. A
RT   genetically distinct family of HLA-A antigens comprising A29, A31,
RT   A32, and Aw33, but probably not A30.";
RL   J. Immunol. 143:3371-3378(1989).
RN   [2]
RP   SEQUENCE FROM N.A. (A*3002).
RX   MEDLINE=93056508; PubMed=1431115;
RA   Madrigal J.A., Belich M.P., Hildebrand W.H., Benjamin R.J.,
RA   Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L.,
RA   Martell R.W., du Toit E.D., Parham P.;
RT   "Distinctive HLA-A,B antigens of black populations formed by
RT   interallelic conversion.";
RL   J. Immunol. 149:3411-3415(1992).
RN   [3]
RP   SEQUENCE OF 25-279 FROM N.A. (A*3003).
RX   MEDLINE=93209813; PubMed=8458735;
RA   Choo S.Y., Starling G.C., Anasetti C., Hansen J.A.;
RT   "Selection of an unrelated donor for marrow transplantation
RT   facilitated by the molecular characterization of a novel HLA-A
RT   allele.";
RL   Hum. Immunol. 36:20-26(1993).
RN   [4]
RP   SEQUENCE FROM N.A. (A*3001).
RX   MEDLINE=95176329; PubMed=7871528;
RA   Olerup O., Daniels T.J., Baxter-Lowe L.;
RT   "Correct sequence of the A*3001 allele obtained by PCR-SSP typing and
RT   automated nucleotide sequencing.";
RL   Tissue Antigens 44:265-267(1994).
RN   [5]
RP   SEQUENCE FROM N.A. (A*3004).
RA   Krausa P., Carcassi C., Orru S., Bodmer J.G., Browning M.J.,
RA   Contu L.;
RL   Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A. (A*300X).
RA   Cox S.T., Mcwhinnie A.J., Madrigal A.J., Little A.M.;
RT   "New A*30 HLA allele found in an Afro-Caribbean bone marrow donor.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 26-206 FROM N.A. (A*3004).
RA   Lienert K., Gao X., McCluskey J.;
RL   Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE OF 28-205 FROM N.A. (A*3004).
RX   MEDLINE=96124443; PubMed=8560452;
RA   Blasczyk R., Wehling J., Paessler M., Hahn U., Huhn D., Salama A.;
RT   "A novel HLA-A30 allele (A*3004) identified by single-strand
RT   conformation polymorphism analysis and confirmed by solid-phase
RT   sequencing.";
RL   Tissue Antigens 46:322-326(1995).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-30 ARE KNOWN: A*3001
CC       (A30.3), A*3002, A*3003 AND A*3004. THE SEQUENCE SHOWN IS THAT OF
CC       A*3001.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30576; AAA59612.1; -.
DR   EMBL; X61702; CAA43871.1; -.
DR   EMBL; M93657; AAA58650.1; -.
DR   EMBL; U07234; AAA70162.1; -.
DR   EMBL; Z34921; CAA84401.1; -.
DR   EMBL; U19734; AAB53658.1; -.
DR   EMBL; U18988; AAB53658.1; JOINED.
```

```
DR   EMBL; U24261; AAB50434.1; -.
DR   EMBL; AJ249308; CAB57306.1; -.
DR   EMBL; AJ249309; CAB57306.1; JOINED.
DR   EMBL; AJ249310; CAB57306.1; JOINED.
DR   EMBL; AJ249311; CAB57306.1; JOINED.
DR   EMBL; AJ249312; CAB57306.1; JOINED.
DR   EMBL; AJ249313; CAB57306.1; JOINED.
DR   EMBL; AJ249314; CAB57306.1; JOINED.
DR   EMBL; AJ249315; CAB57306.1; JOINED.
DR   EMBL; X83770; CAA58723.1; -.
DR   EMBL; X83771; CAA58724.1; -.
DR   PIR; S16766; S16766.
DR   HSSP; Q95352; 1HHK.
DR   MIM; 142800; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     24
FT   CHAIN        25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-30(AW-19) ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN      115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN      207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN      299    308       CONNECTING PEPTIDE.
FT   TRANSMEM    309    332
FT   DOMAIN      333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD    110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID    125    188       BY SIMILARITY.
FT   DISULFID    227    283       BY SIMILARITY.
FT   VARIANT      21     21       H -> Q (IN A*3002, A*3004 AND A*300X).
FT                                 /FTId=VAR_004366.
FT   VARIANT      33     33       S -> Y (IN A*300X).
FT                                 /FTId=VAR_010284.
FT   VARIANT      80     80       R -> G (IN A*3003).
FT                                 /FTId=VAR_004367.
FT   VARIANT      94     94       Q -> H (IN A*3002, A*3003 AND A*3004).
FT                                 /FTId=VAR_004368.
FT   VARIANT     100    101       VD -> EN (IN A*3002, A*3003 AND A*3004).
FT                                 /FTId=VAR_004369.
FT   VARIANT     175    176       RW -> HV (IN A*3004).
FT                                 /FTId=VAR_004370.
FT   VARIANT     176    176       W -> R (IN A*3002 AND A*3003).
FT                                 /FTId=VAR_004371.
FT   VARIANT     180    180       L -> W (IN A*3004).
FT                                 /FTId=VAR_004372.
FT   CONFLICT     33     33       S -> F (IN REF. 1).
SQ   SEQUENCE   365 AA;  40904 MW;  521166D95FB1DC28 CRC64;

  Query Match             33.2%;  Score 505;  DB 1;  Length 365;
  Best Local Similarity   39.0%;  Pred. No. 5.8e-36;
  Matches  108;  Conservative   48;  Mismatches  113;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: |     |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFSTSVSRPGSGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-RPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | ::::|       ||  |:   :| |:  |||:|:: ||::  |     || :: |||
Db      85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYEQHAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  |: :||  :  | | |||||  |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITQRKWEAAR-WAEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   :||| ::    |||| || :|| ||: | :| :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |||||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 14
1A74_HUMAN
ID   1A74_HUMAN     STANDARD;      PRT;   365 AA.
AC   P30459; O46874;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   HLA class I histocompatibility antigen, A-74 alpha chain precursor
DE   (Aw-74) (Aw-19).
GN   HLA-A OR HLAA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (A*7401).
RX   MEDLINE=93056508; PubMed=1431115;
RA   Madrigal J.A., Belich M.P., Hildebrand W.H., Benjamin R.J.,
RA   Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L.,
RA   Martell R.W., du Toit E.D., Parham P.;
RT   "Distinctive HLA-A,B antigens of black populations formed by
RT   interallelic conversion.";
RL   J. Immunol. 149:3411-3415(1992).
RN   [2]
RP   SEQUENCE FROM N.A. (A*7401).
RA   Hurley C.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-206 FROM N.A. (A*7402).
RC   TISSUE=Blood;
RA   Blasczyk R.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   -!- POLYMORPHISM: THE FOLLOWING ALLELES OF A-74 ARE KNOWN: A*7401
CC       AND A*7402. THE SEQUENCE SHOWN IS THAT OF A*7401.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61701; CAA43870.1; -.
DR   EMBL; U17569; AAA56779.1; -.
DR   EMBL; U17570; AAA56780.1; -.
DR   EMBL; AJ223060; CAA11063.1; -.
DR   PIR; S16770; S16770.
DR   HSSP; Q95352; 1HHK.
DR   MIM; 142800; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT   SIGNAL        1     24
FT   CHAIN        25    365       HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-74 ALPHA CHAIN.
FT   DOMAIN       25    114       EXTRACELLULAR ALPHA-1.
```

```
FT   DOMAIN       115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN       207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN       299    308       CONNECTING PEPTIDE.
FT   TRANSMEM     309    332
FT   DOMAIN       333    365       CYTOPLASMIC TAIL.
FT   CARBOHYD     110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   DISULFID     125    188       BY SIMILARITY.
FT   DISULFID     227    283       BY SIMILARITY.
FT   VARIANT       23     23       R -> W (IN A*7402).
FT                                 /FTId=VAR_010375.
SQ   SEQUENCE   365 AA;  40891 MW;  C94BF2569A852371 CRC64;

  Query Match             33.2%;  Score 505;  DB 1;  Length 365;
  Best Local Similarity   39.0%;  Pred. No. 5.8e-36;
  Matches  108;  Conservative   50;  Mismatches  111;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | ::::|       ||  |:   :| |:  |||:|:: ||::  |     || :  |||
Db      85 DQETRNVKAHSQTDRVDLGTLRGYYNQSEAGSHTIQMMYGCDVGPDGRLLRGYQQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  |: :||  :: | | |||||  |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITQRKWEAARV-AEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| | :: || |:|||||| |:| || :|| : ||
Db     263 GTFQKWASVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 15
1A03_PANTR
ID   1A03_PANTR     STANDARD;      PRT;   365 AA.
AC   P13748;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   CHLA class I histocompatibility antigen, A-108 alpha chain precursor.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89030641; PubMed=2460344;
RA   Mayer W.E., Jonker M., Klein D., Ivanyi P., van Seventer G.,
RA   Klein J.;
RT   "Nucleotide sequences of chimpanzee MHC class I alleles: evidence for
RT   trans-species mode of evolution.";
RL   EMBO J. 7:2765-2774(1988).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC       THE IMMUNE SYSTEM.
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X13113; CAA31505.1; ALT_INIT.
DR   PIR; S03535; S03535.
DR   HSSP; O19673; 1HSB.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   MHC I; Transmembrane; Glycoprotein; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    365       CHLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT                                 A-108 ALPHA CHAIN.
FT   DOMAIN        25    114       EXTRACELLULAR ALPHA-1.
FT   DOMAIN       115    206       EXTRACELLULAR ALPHA-2.
FT   DOMAIN       207    298       EXTRACELLULAR ALPHA-3.
FT   DOMAIN       299    308       CONNECTING PEPTIDE.
FT   TRANSMEM     309    332
FT   DOMAIN       333    365       CYTOPLASMIC TAIL.
FT   DISULFID     125    188       BY SIMILARITY.
FT   DISULFID     227    283       BY SIMILARITY.
FT   CARBOHYD     110    110       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   365 AA;  40822 MW;  48CC757055221FC3 CRC64;

  Query Match             33.0%;  Score 503;  DB 1;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 8.5e-36;
  Matches  109;  Conservative   46;  Mismatches  114;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | :   |     |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | ::::|       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db      85 DQETRNMKASAQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||   |  || :||     | | |||||  |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAAAMAAQITKRKWEAAH-AAEQLRAYLEGRCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAVITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || ||||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


Search completed: March 31, 2003, 14:07:23
Job time : 12 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 31, 2003, 14:07:04 ; Search time 19 Seconds
(without alignments)
1396.479 Million cell updates/sec

Title: US-10-092-404-1
Perfect score: 1522
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1149	75.5	359	2	JC5382	hereditary hemochr
2	541.5	35.6	341	2	A57136	class I histocompa
3	522	34.3	361	1	HLRB	MHC class I histoc
4	522	34.3	361	2	I46858	MHC class I RLA pr
5	519	34.1	332	2	S06424	MHC class I histoc
6	516	33.9	365	2	I36961	MHC class I protei
7	515	33.8	361	2	B27638	MHC class I histoc
8	514	33.8	365	2	I83063	A11.2 - human
9	513	33.7	365	2	A47636	MHC class I histoc
10	513	33.7	365	2	I56039	HLA-A30.3 precurso
11	511	33.6	370	1	HLHUA3	MHC class I histoc
12	509	33.4	365	2	I38439	MHC class I histoc
13	508	33.4	365	2	I37542	MHC class I histoc
14	508	33.4	365	2	I38442	gene HLA-A-0205 pr
15	508	33.4	365	2	I61902	MHC class I histoc
16	507	33.3	365	2	I72170	MHC class I histoc
17	507	33.3	365	2	I38441	gene HLA-A-6802 pr
18	505	33.2	365	1	HLHUA2	MHC class I histoc
19	505	33.2	365	2	I37482	MHC class I histoc
20	505	33.2	365	2	I38519	MHC class I histoc
21	505	33.2	365	2	I84448	MHC class I histoc
22	504	33.1	365	2	I38610	MHC class I histoc
23	504	33.1	365	2	I37470	HLA-A*0210 - human
24	503	33.0	355	2	T28149	MHC class I histoc
25	503	33.0	364	2	S03535	class I histocompa
26	502	33.0	365	2	I37476	MHC class I histoc
27	502	33.0	365	2	I37478	MHC class I histoc
28	502	33.0	365	2	I38443	gene HLA-A-0203 pr
29	502	33.0	365	2	I61857	MHC HLA-A2.4a chai
30	501.5	33.0	341	2	JC5663	major histocompati
31	501	32.9	357	2	I36965	MHC class I protei
32	500.5	32.9	362	2	A45845	MHC class I histoc
33	500	32.9	365	2	I61856	MHC class I histoc
34	500	32.9	365	2	I54493	MHC class I histoc
35	499	32.8	273	1	HLHU69	MHC class I histoc
36	499	32.8	365	2	S77963	MHC class I histoc
37	499	32.8	365	2	S01171	class I histocompa
38	499	32.8	365	2	I54416	HLA-AW24 protein -
39	498	32.7	365	2	I37483	HLA-Aw34.2 antigen
40	497	32.7	273	1	HLHUAW	MHC class I histoc
41	497	32.7	360	2	A27638	MHC class I histoc
42	497	32.7	365	2	I72171	HLA-Aw33.1, HLA-Aw
43	496.5	32.6	339	2	I56071	MHC class I histoc
44	496	32.6	279	2	JX0353	zinc-alpha 2-glyco
45	496	32.6	362	2	I68724	MHC class I histoc

ALIGNMENTS

RESULT 1
JC5382
hereditary hemochromatosis protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999
C;Accession: JC5382
R;Hashimoto, K.; Hirai, M.; Kurosawa, Y.
Biochem. Biophys. Res. Commun. 230, 35-39, 1997
A;Title: Identification of a mouse homolog for the human hereditary haemochromatosis
A;Reference number: JC5382; MUID:97148566; PMID:9020055
A;Accession: JC5382
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-359 <HAS>
A;Cross-references: GB:U66849; NID:g1519484; PIDN:AAB07525.1; PID:g1519485
C;Comment: This protein plays a role in iron metabolism.
C;Genetics:
A;Gene: mr2
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-359/Product: hereditary haemochromatosis protein #status predicted <MAT>
F;30-117/Domain: alpha 1 #status predicted <ALF1>
F;118-217/Domain: alpha 2 #status predicted <ALF2>
F;218-309/Domain: alpha 3 #status predicted <ALF3>
F;314-340/Domain: transmembrane #status predicted <TRM>
F;341-359/Domain: intracellular #status predicted <INT>

Query Match 75.5%; Score 1149; DB 2; Length 359;
Best Local Similarity 72.6%; Pred. No. 8.1e-88;
Matches 204; Conservative 30; Mismatches 39; Indels 8; Gaps 1;

```
Qy       4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
           ||||| |||||||| |||| |||| ||||||||| |:||||| ||| ||:  : |||:||
Db      30 RSHSLRYLFMGASEPDLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWL 89

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSK--------ESHTLQVILGCEMQEDNSTEGYWK 115
           |||||||||:|| ||||||| |:|||         ||| |||:||||: ||||| |:|:
Db      90 HLSQSLKGWDYMFIVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWR 149

Qy     116 YGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLEL 175
           |||||||||||| ||:| |||| || ||:||: |||||:||| |||:||| ||::||||
Db     150 YGYDGQDHLEFCPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLEL 209

Qy     176 GRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVL 235
           ||||| |||| ||||| |  |: |:|||:||:::|||||:|||| ||:|||:  |: ||
Db     210 GRGVLGQQVPTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVL 269

Qy     236 PNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||| |||||:|||| ||:| |:|||||||||||||   ||
Db     270 PNGDETYQGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWE 310
```

RESULT 2
A57136
class I histocompatibility antigen related protein MR1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 23-Jul-1999
C;Accession: A57136
R;Hashimoto, K.; Hirai, M.; Kurosawa, Y.
Science 269, 693-695, 1995
A;Title: A gene outside the human MHC related to classical HLA class I genes.
A;Reference number: A57136; MUID:95350662; PMID:7624800
A;Accession: A57136
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-341 <HAS>
A;Cross-references: GB:U22963; NID:g940353; PIDN:AAC50174.1; PID:g940354
C;Genetics:
A;Gene: GDB:HLALS
A;Cross-references: GDB:683188; OMIM:600764
A;Map position: 1q25.3-1q25.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 35.6%; Score 541.5; DB 2; Length 341;
Best Local Similarity 39.5%; Pred. No. 1.9e-37;
Matches 107; Conservative 50; Mismatches 111; Indels 3; Gaps 3;

```
Qy       4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
           |:||| |  :| |:   |:  | ::|||||      ||  :|: ||| ||::  ::   |
Db      23 RTHSLRYFRLGVSDPIHGVPEFISVGYVDSHPITTYDSVTRQKEPRAPWMAENLAPDHWE 82

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
           : :| |:||  || |:   :  ::|||  ||| | ::|||: || || |: :| |||||
Db      83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy     124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
           | |  ||| | | :  |   |  || ::      : :|| :| | |:: || |:  | :
Db     142 LIFNKDTLSWLAVDNVAHTIKQAWEANQHELLYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy     184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
            ||||:|    |   || | |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db     202 EPPLVRVNRKETFPGVTALFCKAHGFYPPEIYMTWMKNGEEI-VQEIDYGDILPSGDGTY 260

Qy     243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
           | | :: : |     |:| |||  |:   | |
Db     261 QAWASIELDPQSSNLYSCHVEHCGVHMVLQV 291
```

RESULT 3
HLRB
MHC class I histocompatibility antigen RLA alpha chain precursor (RL-5) - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 22-Jun-1999
C;Accession: A02193
R;Tykocinski, M.L.; Marche, P.N.; Max, E.E.; Kindt, T.J.
J. Immunol. 133, 2261-2269, 1984
A;Title: Rabbit class I MHC genes: cDNA clones define full-length transcripts of an expr
A;Reference number: A02193; MUID:84290724; PMID:6432910
A;Accession: A02193
A;Molecule type: mRNA
A;Residues: 1-361 <TYK>
A;Cross-references: GB:K02441; NID:g1293894; PIDN:AAA98729.1; PID:g165496
A;Note: the source of this protein is a T-lymphoid cell line (RL-5), which has been tran
C;Comment: In contrast to the many antigens expressed in mouse (K, D, and L) and human (
 MHC may therefore differ from the HLA and H-2 loci in having limited complexity.
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterodimer; transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-361/Product: class I histocompatibility antigen RLA alpha chain #status predicted <
F;25-307/Domain: extracellular #status predicted <EXT>
F;25-114/Domain: alpha-1 <EX1>
F;115-206/Domain: alpha-2 <EX2>
F;220-285/Domain: immunoglobulin homology <IMM>
F;308-329/Domain: transmembrane #status predicted <TMM>
F;330-361/Domain: intracellular #status predicted <INT>
F;110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;125-188,227-283/Disulfide bonds: #status predicted

Query Match 34.3%; Score 522; DB 1; Length 361;
Best Local Similarity 40.1%; Pred. No. 8.6e-36;
Matches 111; Conservative 44; Mismatches 114; Indels 8; Gaps 7;

```
Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | :   |   ||   |  :|||||  || :| :  | |:| | ||:  ::  : |
Db      26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db      85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           |::    |   | ||:  |  |: :||     | ::|||||:|   |:: ||:|:  |
Db     145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy     181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   |||| |    |||| || :||  |:: | :| :    :: |  :  | ||
Db     204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || ||||||||:|:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299
```

RESULT 4
I46858
MHC class I RLA precursor - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 21-Jan-2000
C;Accession: I46858
R;Marche, P.N.; Tykocinski, M.L.; Max, E.E.; Kindt, T.J.
Immunogenetics 21, 71-82, 1985
A;Title: Structure of a functional rabbit class I MHC gene: Similarity to human class
A;Reference number: I46858; MUID:85103547; PMID:3917974
A;Accession: I46858
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-361 <MAR>
A;Cross-references: GB:K02819; NID:g165497; PIDN:AAA98730.1; PID:g165498
C;Genetics:
A;Introns: 25/1; 115/1; 207/1; 299/1; 337/1; 348/1
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

Query Match 34.3%; Score 522; DB 2; Length 361;
Best Local Similarity 40.1%; Pred. No. 8.6e-36;
Matches 111; Conservative 44; Mismatches 114; Indels 8; Gaps 7;

```
Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | :   |   ||   |  :|||||  || :| :  | |:| | ||:  ::  : |
Db      26 SHSMRYFYTSVSRPGLGEPRFIIVGYVDDTQFVRFDSDAASPRMEQRAPWM-GQVEPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | :|  |     | |:  | :  :| |   ||| | : |||:  |     || :| |||
Db      85 DQQTQIAKDTAQTFRVNLNTALRYYNQSAAGSHTFQTMFGCEVWADGRFFHGYRQYAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           |::    |   | ||:  |  |: :||     | ::|||||:|   |:: ||:|:  |
Db     145 ADYIALNEDLRSWTAADTAAQNTQRKWEAAG-EAERHRAYLERECVEWLRRYLEMGKETL 203

Qy     181 DQQVPPLVKVTHHVTSS-VTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   |||| |    |||| || :||  |:: | :| :    :: |  :  | ||
Db     204 QRADPPKAHVTHHPASDREATLRCWALGFYPAEISLTWQRDGED-QTQDTELVETRPGGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || ||||||||:|:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCRVQHEGLPEPLTLTWE 299
```

RESULT 5
S06424
MHC class I histocompatibility antigen Ch25 alpha chain precursor - chimpanzee
N;Alternate names: MHC ChlA chain
C;Species: Pan troglodytes (chimpanzee)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 23-Jul-1999
C;Accession: S06424; I36959
R;Lawlor, D.A.; Ward, F.E.; Ennis, P.D.; Jackson, A.P.; Parham, P.
Nature 335, 268-271, 1988
A;Title: HLA-A and B polymorphisms predate the divergence of humans and chimpanzees.
A;Reference number: S06424; MUID:88319000; PMID:3412487
A;Accession: S06424
A;Molecule type: mRNA
A;Residues: 1-332 <LAW>
R;Parham, P.; Lawlor, D.A.; Lomen, C.E.; Ennis, P.D.
J. Immunol. 142, 3937-3950, 1989
A;Title: Diversity and diversification of HLA-A,B,C alleles.
A;Reference number: I36956; MUID:89235215; PMID:2715640
A;Accession: I36959
A;Molecule type: mRNA
A;Residues: 1-332 <RES>
A;Cross-references: GB:M24047; NID:g176818; PIDN:AAA35426.1; PID:g553155
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: glycoprotein; membrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-114/Domain: alpha-1 #status predicted <EX1>
F;115-206/Domain: alpha-2 #status predicted <EX2>
F;220-285/Domain: immunoglobulin homology <IMM>
F;307-331/Domain: transmembrane #status predicted <TMM>
F;110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;125-188,227-283/Disulfide bonds: #status predicted

Query Match 34.1%; Score 519; DB 2; Length 332;
Best Local Similarity 40.1%; Pred. No. 1.4e-35;
Matches 111; Conservative 44; Mismatches 114; Indels 8; Gaps 7;

```
Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | ::| |       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db      85 DQETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::   |  | ||:  |  || :||     | | ||||| |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |||||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299
```

RESULT 6
I36961
MHC class I protein - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C;Accession: I36961
R;Lawlor, D.A.; Warren, E.; Ward, F.E.; Parham, P.
Immunol. Rev. 113, 147-185, 1990
A;Title: Comparison of class I MHC alleles in humans and apes.
A;Reference number: I36961; MUID:90201944; PMID:1690682
A;Accession: I36961
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M30678; NID:g176822; PIDN:AAA87970.1; PID:g176823
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

Query Match 33.9%; Score 516; DB 2; Length 365;
Best Local Similarity 39.7%; Pred. No. 2.7e-35;
Matches 110; Conservative 45; Mismatches 114; Indels 8; Gaps 7;

```
Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           : ::| |       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db      85 DEETRSAKAHSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::   |  | ||:  |  || :||     | | ||||| |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQRRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |||||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299
```

RESULT 7
B27638
MHC class I histocompatibility antigen alpha chain precursor (BL3-7) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 16-Feb-1997
C;Accession: B27638
R;Ennis, P.D.; Jackson, A.P.; Parham, P.
J. Immunol. 141, 642-651, 1988
A;Title: Molecular cloning of bovine class I MHC cDNA.
A;Reference number: A92826; MUID:88258075; PMID:3133413
A;Accession: B27638
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-361 <ENN>
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: heterodimer; transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-361/Product: MHC class I histocompatibility antigen, BoLA alpha chain (BL3-7) #s
F;220-285/Domain: immunoglobulin homology <IMM>

Query Match 33.8%; Score 515; DB 2; Length 361;
Best Local Similarity 38.9%; Pred. No. 3.3e-35;
Matches 109; Conservative 50; Mismatches 113; Indels 8; Gaps 7;

```
Qy       2 LLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISS 59
           |  |||| | : | |   ||   | |:|||||  || :| :  : | ||| ||:
Db      23 LAGSHSLRYFYTGVSRPGLGEPRFIAVGYVDDTQFVRFDSDAPNPREEPRVPWMEQE-GP 81

Qy      60 QMWLQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYG 117
           : | : ::  |    :| ||  |:   :| |:  || :| : ||::  |     |:|::|
Db      82 EYWDRNTRIYKDTAQIFRVDLNTLRGYYNQSETGSHNIQAMYGCDVGPDGRLLRGFWQFG 141

Qy     118 YDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGR 177
           |||:|::    :   | ||:  |  || :||     |   | ||| :|   |:: || |:
Db     142 YDGRDYIALNEELRSWTAADTAAQITKRKWEAAG-AAETWRNYLEGECVEWLRRYLENGK 200

Qy     178 GVLDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLP 236
           | :  ||   |||| ::    |||| || :||: |:: | :: :    :: |  :  |
Db     201 DTLLRADPPKAHVTHHSISDREVTLRCWALGFYPEEISLTWQREGED-QTQDMELVETRP 259

Qy     237 NGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           :||||:| |  | || ||||||||:|:| || :|| : ||
Db     260 SGDGTFQKWAALVVPSGEEQRYTCRVQHEGLQEPLTLRWE 299
```

RESULT 8

```
I83063
A11.2 - human
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C;Accession: I83063
R;Lin, L.; Tokunaga, K.; Ishikawa, Y.; Bannai, M.; Kashiwase, K.; Kuwata, S.; Akaza, T.;
Tissue Antigens 43, 78-82, 1994
A;Title: Sequence analysis of serological HLA-A11 split antigens, A11.1 and A11.2.
A;Reference number: I60129; MUID:94287401; PMID:8016845
A;Accession: I83063
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:D16842; NID:g540517; PIDN:BAA04118.1; PID:g487911
C;Genetics:
A;Gene: A1102
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.8%;  Score 514;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 4e-35;
  Matches  109;  Conservative   47;  Mismatches  113;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | :   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFYTSVSRPGRGKPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | ::::|       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db      85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   :||| ::    |||| || :||  ||: | :|  :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |||||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 9
A47636
MHC class I histocompatibility antigen HLA-A11 alpha chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 28-Apr-1995 #text_change 23-Jul-1999
C;Accession: S03536; S03694; A47636; I60129
R;Mayer, W.E.; Jonker, M.; Klein, D.; Ivanyi, P.; van Seventer, G.; Klein, J.
EMBO J. 7, 2765-2774, 1988
A;Title: Nucleotide sequences of chimpanzee MHC class I alleles: evidence for trans-spec
A;Reference number: S01171; MUID:89030641; PMID:2460344
A;Accession: S03536
A;Molecule type: mRNA
A;Residues: 1-365 <MAY>
A;Cross-references: EMBL:X13111; NID:g32138; PIDN:CAA31503.1; PID:g32139
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
A;Accession: S03694
A;Molecule type: mRNA
A;Residues: 1-42,'K',44-298 <MA2>
A;Cross-references: EMBL:X13112; NID:g32142; PIDN:CAA31504.1; PID:g32143
A;Note: this allele is designated A*1102 (formerly A11K, A11.2)
R;Cowan, E.P.; Jelachich, M.L.; Biddison, W.E.; Coligan, J.E.
Immunogenetics 25, 241-250, 1987
A;Title: DNA sequence of HLA-A11: remarkable homology with HLA-A3 allows identification
A;Reference number: A47636; MUID:87192928; PMID:2437024
A;Accession: A47636
A;Molecule type: DNA
A;Residues: 26-365 <COW>
A;Cross-references: GB:M16007; GB:M16008; GB:M16009; GB:M16010; NID:g184130; PIDN:AAA654
A;Note: the authors translated the codon GAC for residue 89 as Ala, CCG for residue 104
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
R;Lin, L.; Tokunaga, K.; Ishikawa, Y.; Bannai, M.; Kashiwase, K.; Kuwata, S.; Akaza,
Tissue Antigens 43, 78-82, 1994
A;Title: Sequence analysis of serological HLA-A11 split antigens, A11.1 and A11.2.
A;Reference number: I60129; MUID:94287401; PMID:8016845
A;Accession: I60129
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:D16841; NID:g540516; PIDN:BAA04117.1; PID:g487909
A;Note: this allele is designated A*1101 (formerly A11E, A11.1)
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-365/Product: class I histocompatibility antigen alpha chain #status predicted <M
F;25-298/Domain: extracellular #status predicted <EXT>
F;220-285/Domain: immunoglobulin homology <IMM>
F;299-337/Domain: transmembrane #status predicted <TMM>
F;338-365/Domain: intracellular #status predicted <INT>

  Query Match             33.7%;  Score 513;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 4.8e-35;
  Matches  109;  Conservative   47;  Mismatches  113;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | :   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | ::::|       ||  |:   :| |:: |||:|:: ||::  |     || :  |||
Db      85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEDGSHTIQIMYGCDVGPDGRFLRGYRQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||     | | |||||  |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITKRKWEAAH-AAEQQRAYLEGRCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   :||| ::    |||| || :||  ||: | :|  :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |||||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 10
I56039
HLA-A30.3 precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C;Accession: I56039
R;Kato, K.; Trapani, J.A.; Allopenna, J.; Dupont, B.; Yang, S.Y.
J. Immunol. 143, 3371-3378, 1989
A;Title: Molecular analysis of the serologically defined HLA-Aw19 antigens. A genetic
A;Reference number: I56039; MUID:90038496; PMID:2478623
A;Accession: I56039
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M30576; NID:g187646; PIDN:AAA59612.1; PID:g386878
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.7%;  Score 513;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 4.8e-35;
  Matches  109;  Conservative   48;  Mismatches  112;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
```

```
Db      26 SHSMRYFFTSVSRPGSGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-RPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | ::::|       ||  |:   :| |:  |||:|:: ||::  |     || :: |||
Db      85 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYEQHAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  |: :||  :  | | |||||  |   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITQRKWEAAR-WAEQLRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |||||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 11
HLHUA3
MHC class I histocompatibility antigen HLA-A3 alpha chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 02-Sep-1997
C;Accession: A02192
R;Strachan, T.; Sodoyer, R.; Damotte, M.; Jordan, B.R.
EMBO J. 3, 887-894, 1984
A;Title: Complete nucleotide sequence of a functional class I HLA gene, HLA-A3: implicat
A;Reference number: A02192; MUID:84207948; PMID:6609814
A;Accession: A02192
A;Molecule type: DNA
A;Residues: 1-370 <STR>
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
A;Introns: 30/1; 120/1; 212/1; 304/1; 343/1; 354/1; 370/1
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplantati
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-370/Product: class I histocompatibility antigen HLA-A3 alpha chain #status predicte
F;30-312/Domain: extracellular #status predicted <EXT>
F;30-119/Domain: alpha-1 <EX1>
F;120-211/Domain: alpha-2 <EX2>
F;225-290/Domain: immunoglobulin homology <IMM>
F;313-337/Domain: transmembrane #status predicted <TMM>
F;338-370/Domain: intracellular #status predicted <INT>
F;115/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;232-288/Disulfide bonds: #status predicted

  Query Match             33.6%;  Score 511;  DB 1;  Length 370;
  Best Local Similarity   39.6%;  Pred. No. 7.2e-35;
  Matches  110;  Conservative   47;  Mismatches  111;  Indels   10;  Gaps    8;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      31 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 89

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           | ::::|       ||  |:   :| |:  |||:|:: ||::  |     || :  |||
Db      90 DQETRNVKAQSQTDRVDLGTLRGYYNQSEAGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 149

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWE-RHKIRARQNRAYLERDCPAQLQQLLELGRGV 179
           :|::    |   | ||:  |  || :||  |:  | | ||||:  |   |:: || |:
Db     150 KDYIALNEDLRSWTAADMAAQITKRKWEAAHE--AEQLRAYLDGTCVEWLRRYLENGKET 207

Qy     180 LDQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNG 238
           | :  ||   :||| ::    |||| || :||  ||: | :| :    :: |  :  | |
Db     208 LQRTDPPKTHMTHHPISDHEATLRCWALGFYPAEITLTWQRDGED-QTQDTELVETRPAG 266

Qy     239 DGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |||:| |  : || |||||||| |:| || :|| : ||
Db     267 DGTFQKWAAVVVPSGEEQRYTCHVQHEGLPKPLTLRWE 304


RESULT 12
I38439
MHC class I histocompatibility antigen HLA-A*8001 precursor - human
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 21-Jan-2000
C;Accession: I59638; I38439
R;Domena, J.D.; Hildebrand, W.H.; Bias, W.B.; Parham, P.
Tissue Antigens 42, 156-159, 1993
A;Title: A sixth family of HLA-A alleles defined by HLA-A*8001.
A;Reference number: I59638; MUID:94112691; PMID:8284791
A;Accession: I59638
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <DOM>
A;Cross-references: GB:L18898; NID:g306853; PIDN:AAA17012.1; PID:g306854
R;Balas, A.; Garcia-Sanchez, F.; Gomez-Reino, F.; Vicario, J.L.
Immunogenetics 39, 452, 1994
A;Title: Characterization of a new and highly distinguishable HLA-A allele in a Spani
A;Reference number: I38439; MUID:94245293; PMID:8188325
A;Accession: I38439
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <BAL>
A;Cross-references: EMBL:U03754; NID:g432407; PIDN:AAC04322.1; PID:g432408
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.4%;  Score 509;  DB 2;  Length 365;
  Best Local Similarity   38.3%;  Pred. No. 1e-34;
  Matches  106;  Conservative   53;  Mismatches  110;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDSQFVQFDSDAASQRMEPRAPWIEQE-EPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQEDNS-TEGYWKYGYDG 120
           : ::::|       :  |:   :| |:: |||:|:: ||::  |     || :  |||
Db      85 DEETRNVKAHSQTNRANLGTLRGYYNQSEDGSHTIQIMYGCDVGSDGRFLRGYRQDAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||  : || | ||||| :|   |:: || |:  |
Db     145 KDYIALNEDLRSWTAADMAAQITKRKWEAAR-RAEQLRAYLEGECVDGLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
           :  ||   :||| ::    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     204 QRTDPPKTHMTHHPISDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |:|:|||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGKEKRYTCHVQHEGLPEPLTLRWE 299


RESULT 13
I37542
MHC class I histocompatibility antigen HLA-A2 alpha chain (allele A*0216) precursor -
C;Species: Homo sapiens (man)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C;Accession: I37542; S49582
R;Barouch, D.; Krausa, P.; Bodmer, J.; Browning, M.J.; McMichael, A.J.
Immunogenetics 41, 388, 1995
A;Title: Identification of a novel HLA-A2 subtype, HLA-A*0216.
A;Reference number: I37542; MUID:95278976; PMID:7759139
A;Accession: I37542
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: EMBL:Z46633; NID:g575248; PIDN:CAA86602.1; PID:g575249
```

```
A;Note: submitted to the EMBL Data Library, November 1994
C;Genetics:
A;Gene: hla-A
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.4%;  Score 508;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 1.3e-34;
  Matches  109;  Conservative   46;  Mismatches  114;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
           :: :|      ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db      85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||   : | | |||||  :|   |:: || |:  |
Db     145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQLRAYLEGECVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |:|||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 14
I38442
gene HLA-A-0205 protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C;Accession: I38442
R;Holmes, N.; Ennis, P.; Wan, A.M.; Denney, D.W.; Parham, P.
J. Immunol. 139, 936-941, 1987
A;Title: Multiple genetic mechanisms have contributed to the generation of the HLA-A2/A2
A;Reference number: I38441; MUID:87252273; PMID:3496393
A;Accession: I38442
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-365 <RES>
A;Cross-references: EMBL:U03862; NID:g432436; PIDN:AAA03603.1; PID:g432437
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.4%;  Score 508;  DB 2;  Length 365;
  Best Local Similarity   39.7%;  Pred. No. 1.3e-34;
  Matches  110;  Conservative   44;  Mismatches  115;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | :   |    |   | |:|||||  || :| :  |||:||| ||:      : |
Db      26 SHSMRYFYTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASRRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
             :: :|      ||  |:   :| |:  ||||| : ||::  |     || :| |||
Db      85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTLQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db     145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQWRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |:|||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


RESULT 15
I61902
MHC class I histocompatibility antigen HLA-A alpha chain precursor - human (isolate A
C;Species: Homo sapiens (man)
A;Variety: isolate A*0212
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 23-Jul-1999
C;Accession: I61902
R;Belich, M.P.; Madrigal, J.A.; Hildebrand, W.H.; Zemmour, J.; Williams, R.C.; Luz, R
Nature 357, 326-329, 1992
A;Title: Unusual HLA-B alleles in two tribes of Brazilian Indians.
A;Reference number: I37120; MUID:92269955; PMID:1317015
A;Accession: I61902
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-365 <RES>
A;Cross-references: GB:M84378; NID:g187625; PIDN:AAA59604.1; PID:g187626
A;Experimental source: cell line KRC 033; isolate A*0212
C;Genetics:
A;Gene: GDB:HLA-A
A;Cross-references: GDB:119310; OMIM:142800
A;Map position: 6p21.3-6p21.3
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology
C;Keywords: transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-365/Product: MHC class I histocompatibility antigen HLA-A alpha chain #status pr
F;220-285/Domain: immunoglobulin homology <IMM>

  Query Match             33.4%;  Score 508;  DB 2;  Length 365;
  Best Local Similarity   39.4%;  Pred. No. 1.3e-34;
  Matches  109;  Conservative   45;  Mismatches  115;  Indels    8;  Gaps    7;

Qy       5 SHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHE--SRRVEPRTPWVSSRISSQMW 62
           |||: | |   |    |   | |:|||||  || :| :  |:|:||| ||:      : |
Db      26 SHSMRYFFTSVSRPGRGEPRFIAVGYVDDTQFVRFDSDAASQRMEPRAPWIEQE-GPEYW 84

Qy      63 LQLSQSLKGWDHMFTVDFWTIMENHNHSKE-SHTLQVILGCEMQED-NSTEGYWKYGYDG 120
             :: :|      ||  |:   :| |:  |||:| : ||::  |     || :| |||
Db      85 DGETRKVKAHSQTHRVDLGTLRGYYNQSEAGSHTVQRMYGCDVGSDWRFLRGYHQYAYDG 144

Qy     121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180
           :|::    |   | ||:  |  || :||   : | | |||||  |   |:: || |:  |
Db     145 KDYIALKEDLRSWTAADMAAQTTKHKWEAAHV-AEQQRAYLEGTCVEWLRRYLENGKETL 203

Qy     181 DQQVPPLVKVTHH-VTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGD 239
             :   |   :||| |:    |||| ||::||  ||: | :| :    :: |  :  | ||
Db     204 QRTDAPKTHMTHHAVSDHEATLRCWALSFYPAEITLTWQRDGED-QTQDTELVETRPAGD 262

Qy     240 GTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           ||:| |  : || |:|||||| |:| || :|| : ||
Db     263 GTFQKWAAVVVPSGQEQRYTCHVQHEGLPKPLTLRWE 299


Search completed: March 31, 2003, 14:09:09
Job time : 20 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: March 31, 2003, 14:07:04 ; Search time 33 Seconds
(without alignments)
1723.303 Million cell updates/sec

Title: US-10-092-404-1
Perfect score: 1522
Sequence: 1 RLLRSHSLHYLFMGASEQDL.....RYTCQVEHPGLDQPLIVIWE 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1387	91.1	325	4	Q96KU8	Q96ku8 homo sapien
2	1247	81.9	348	6	Q9GL42	Q9gl42 dicerorhinu
3	1245	81.8	348	6	Q9GKZ0	Q9gkz0 ceratotheri
4	1241	81.5	348	6	Q9GL41	Q9gl41 rhinoceros
5	1238	81.3	348	6	Q9GL43	Q9gl43 diceros bic
6	1233	81.0	322	6	Q9GK81	Q9gk81 diceros bic
7	1149	75.5	359	11	Q9D754	Q9d754 mus musculu
8	952	62.5	256	4	Q96KU7	Q96ku7 homo sapien
9	948	62.3	256	4	Q9HC68	Q9hc68 homo sapien
10	899	59.1	165	4	Q9HC70	Q9hc70 homo sapien
11	859	56.4	242	4	Q9HC64	Q9hc64 homo sapien
12	802	52.7	272	11	Q9R105	Q9r105 rattus norv
13	662	43.5	161	4	Q9HC83	Q9hc83 homo sapien
14	592	38.9	116	4	Q9HC69	Q9hc69 homo sapien
15	542.5	35.6	340	7	Q9BD50	Q9bd50 pongo pygma
16	541.5	35.6	334	7	Q9TQK3	Q9tqk3 homo sapien
17	541.5	35.6	341	4	Q9NPL2	Q9npl2 homo sapien
18	541.5	35.6	341	7	Q9BCU3	Q9bcu3 pan troglod
19	541.5	35.6	341	7	Q95460	Q95460 homo sapien
20	539.5	35.4	354	7	Q95HB3	Q95hb3 anas platyr
21	538.5	35.4	341	7	Q9BCU4	Q9bcu4 pan troglod
22	530	34.8	105	4	Q9HC71	Q9hc71 homo sapien
23	519	34.1	332	7	Q30990	Q30990 pan troglod
24	519	34.1	365	7	Q9TPL7	Q9tpl7 pan troglod
25	517	34.0	365	7	Q9TQP6	Q9tqp6 homo sapien
26	517	34.0	371	7	Q9TQP7	Q9tqp7 homo sapien
27	516	33.9	364	7	O19243	O19243 sus scrofa
28	516	33.9	365	7	Q9MYG4	Q9myg4 homo sapien
29	515	33.8	168	4	Q96KU5	Q96ku5 homo sapien
30	514	33.8	365	7	Q29747	Q29747 homo sapien
31	513	33.7	273	7	Q95IG6	Q95ig6 homo sapien
32	513	33.7	352	7	Q8SPA9	Q8spa9 sus scrofa
33	513	33.7	365	7	Q9MYI5	Q9myi5 homo sapien
34	512	33.6	129	4	Q9UK37	Q9uk37 homo sapien
35	512	33.6	330	7	O19356	O19356 macaca mula
36	512	33.6	331	7	O02944	O02944 macaca mula
37	512	33.6	333	7	O98030	O98030 papio anubi
38	512	33.6	333	7	O98031	O98031 papio anubi
39	511	33.6	330	7	O02946	O02946 macaca mula
40	511	33.6	330	7	O02947	O02947 macaca mula
41	511	33.6	365	7	O19756	O19756 homo sapien
42	510	33.5	331	7	O02945	O02945 macaca mula
43	510	33.5	357	7	Q30886	Q30886 pan paniscu
44	510	33.5	363	7	Q9MWK4	Q9mwk4 gorilla gor
45	510	33.5	363	7	Q9MXI5	Q9mxi5 pan troglod

ALIGNMENTS

RESULT 1
Q96KU8
ID Q96KU8 PRELIMINARY; PRT; 325 AA.
AC Q96KU8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hemochromatosis protein.
GN HFE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Oliva R.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA OIiva R., Sanchez M.;
RT "Identification of different alternative splicing forms of the HFE gene.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AJ249335; CAC67792.1; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Glycoprotein; Transmembrane.
SQ SEQUENCE 325 AA; 37514 MW; 626343ACFAA862EF CRC64;

Query Match 91.1%; Score 1387; DB 4; Length 325;

```
  Best Local Similarity   100.0%;  Pred. No. 2e-126;
  Matches  249;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      28 LGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDFWTIMENH 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 LGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDFWTIMENH 86

Qy      88 NHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEW 147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87 NHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEW 146

Qy     148 ERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTSSVTTLRCRALN 207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147 ERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTSSVTTLRCRALN 206

Qy     208 YYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGL 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     207 YYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGL 266

Qy     268 DQPLIVIWE 276
           |||||||||
Db     267 DQPLIVIWE 275


RESULT 2
Q9GL42
ID   Q9GL42      PRELIMINARY;      PRT;   348 AA.
AC   Q9GL42;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   HFE protein.
OS   Dicerorhinus sumatrensis (Sumatran rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerorhinus.
OX   NCBI_TaxID=89632;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE Polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AY007543; AAG23703.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   348 AA;  39740 MW;  518BFD357AB83B90 CRC64;

  Query Match           81.9%;  Score 1247;  DB 6;  Length 348;
  Best Local Similarity   81.7%;  Pred. No. 8.5e-113;
  Matches  223;  Conservative   20;  Mismatches   30;  Indels    0;  Gaps    0;

Qy       4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
           ||||| |||||||||:| || ||||||||||:||| |:||||| | |  ||     ||:||
Db      26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
           |||||||||||||| |||||||:||||||||||||||||||:|||||| |:||||||||||
Db      86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy     124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
           |||||:|||||||| ||  |||||| :||||:|||||||||||| ||| ||||||||||||
Db     146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy     184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
           |||||||||||| |:||||||:|||:|||||||:||||::|:| |: | |||||:||||||
Db     206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy     244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |: |||||||||||||||||||||||||   ||
Db     266 SWVALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 3
Q9GKZ0
ID   Q9GKZ0      PRELIMINARY;      PRT;   348 AA.
AC   Q9GKZ0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   HFE protein.
OS   Ceratotherium simum (White rhinoceros) (Square-lipped rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
OX   NCBI_TaxID=9807;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE Polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AY007541; AAG23701.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   348 AA;  39822 MW;  2523016ECE9FBE91 CRC64;

  Query Match           81.8%;  Score 1245;  DB 6;  Length 348;
  Best Local Similarity   82.1%;  Pred. No. 1.3e-112;
  Matches  224;  Conservative   18;  Mismatches   31;  Indels    0;  Gaps    0;

Qy       4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
           ||||| |||||||||:| || ||||||||||:||| |:||||| | |  ||     ||:||
Db      26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
           |||||||||||||| |||||||:||||||||||||||||||:|||||| |:||||||||||
Db      86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy     124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
           |||||:|||||||| ||  |||||| :||||:|||||||||||| ||| ||||||||||||
Db     146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy     184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
           |||||||||||| |:||||||:|||:|||||||:||||::|:| |: | |||||:||||||
Db     206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPMDVKDAESKDVLPSGDGTYQ 265

Qy     244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
           |  |||||||||||||||||||||||||   ||
Db     266 SWEALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 4
Q9GL41
```

```
ID   Q9GL41      PRELIMINARY;      PRT;   348 AA.
AC   Q9GL41;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   HFE protein.
OS   Rhinoceros unicornis (Greater Indian rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Rhinoceros.
OX   NCBI_TaxID=9809;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE Polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AY007544; AAG23704.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   348 AA;  39743 MW;  F2723D57A327A6B4 CRC64;

  Query Match             81.5%;  Score 1241;  DB 6;  Length 348;
  Best Local Similarity   81.3%;  Pred. No. 3.3e-112;
  Matches  222;  Conservative   20;  Mismatches   31;  Indels    0;  Gaps    0;

Qy       4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
           ||||| |||||||:| || |||||||||:||| |:||||| | |  ||     ||:||
Db      26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
           |||||||||||| |||||||:|||||||||||||||||:|||||| |:| ||||||||
Db      86 QLSQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWMYGYDGQDH 145

Qy     124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
           |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db     146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy     184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
           |||||||||||  |:|||||:|||:|||||||:||||::|:|  |:  | |||||:||||||
Db     206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy     244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
             |: |||||||||||||||||||||||   ||
Db     266 SWVALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 5
Q9GL43
ID   Q9GL43      PRELIMINARY;      PRT;   348 AA.
AC   Q9GL43;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   HFE protein.
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX   NCBI_TaxID=9805;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE Polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AY007542; AAG23702.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   348 AA;  39818 MW;  4D95E7B01E48FB90 CRC64;

  Query Match             81.3%;  Score 1238;  DB 6;  Length 348;
  Best Local Similarity   81.3%;  Pred. No. 6.4e-112;
  Matches  222;  Conservative   20;  Mismatches   31;  Indels    0;  Gaps    0;

Qy       4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
           ||||| |||||||:| || |||||||||:||| |:||||| | |  ||     ||:||
Db      26 RSHSLRYLFMGASERDHGLPLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWL 85

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
           ||:|||||||||| |||||||:|||||||||||||||||:|||||| |:||||||||||
Db      86 QLTQSLKGWDHMFIVDFWTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDH 145

Qy     124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
           |||||:|||||||| ||  |||||| :||||:|||||||||| ||| ||||||||||||
Db     146 LEFCPETLDWRAAESRALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQ 205

Qy     184 VPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQ 243
           |||||||||||  |:|||||:|||:|||||||:||||::|:|  |:  | |||||:||||||
Db     206 VPPLVKVTHHVASAVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQ 265

Qy     244 GWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
             |  ||||||||||||||||||||||||   ||
Db     266 SWEALAVPPGEEQRYTCQVEHPGLDQPLTATWE 298


RESULT 6
Q9GK81
ID   Q9GK81      PRELIMINARY;      PRT;   322 AA.
AC   Q9GK81;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   HFE (Fragment).
OS   Diceros bicornis (Black rhinoceros).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX   NCBI_TaxID=9805;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West C.J., Worley M., Beutler E.;
RT   "Rhinoceros HFE Polymorphisms.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF301592; AAG39940.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
```

```
Db    121 PQNITMKWLKDKQPMDTKEFEPKDVLPNGDGTYQGWITLAVPP 163


RESULT 11
Q9HC64
ID   Q9HC64      PRELIMINARY;      PRT;   242 AA.
AC   Q9HC64;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hemochromatosis protein splice variant 562-878del.
GN   HFE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20448010; PubMed=11001625;
RA   Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.Y., David V.,
RA   Mosser J.;
RT   "The HFE gene undergoes alternate splicing processes.";
RL   Blood Cells Mol. Dis. 26:155-162(2000).
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF149804; AAG29342.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00129; MHC_I; 1.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   242 AA;  27826 MW;  A021771A22A1F6D8 CRC64;

  Query Match             56.4%;  Score 859;  DB 4;  Length 242;
  Best Local Similarity   60.9%;  Pred. No. 2.5e-75;
  Matches  168;  Conservative    1;  Mismatches    1;  Indels  106;  Gaps    1;

Qy      1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy     61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
          ||||||||||||||||||||||||||||||
Db     83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSK---------------------------- 113

Qy    121 QDHLEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVL 180

Db    114 ------------------------------------------------------------ 113

Qy    181 DQQVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 240
                      ||||||||||:||||||||||||||||||||||||||||||||
Db    114 ----------------VTTLRCRALNHYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDG 156

Qy    241 TYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWE 276
          ||||||||||| ||||||||||||||||||||||||
Db    157 TYQGWITLAVPPEEEQRYTCQVEHPGLDQPLIVIWE 192


RESULT 12
Q9R105
ID   Q9R105      PRELIMINARY;      PRT;   272 AA.
AC   Q9R105;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hemochromatosis gene product HFE splice variant delE2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=TESTIS;
RA   Liew Y.-F., Shaw N.-S.;
RT   "Alternative splice variant of the hemochromatosis gene HFE in iron
RT   overloaded rats.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AF176534; AAD49965.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   272 AA;  30757 MW;  1D91063CCBEF5502 CRC64;

  Query Match             52.7%;  Score 802;  DB 11;  Length 272;
  Best Local Similarity   75.1%;  Pred. No. 1e-69;
  Matches  139;  Conservative   22;  Mismatches   24;  Indels    0;  Gaps    0;

Qy     92 ESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWERHK 151
          ||| |||||||:|||||:|||||||||||||||||:|:| |||||||:|||||:|
Db     39 ESHILQVILGCEVHEDNSTSGFWKYGYDGQDHLEFCPKTLNWSAAEPRAWATKMEWEEHR 98

Qy    152 IRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTSSVTTLRCRALNYYPQ 211
          |||||:| ||:|||| ||:|:||| |||| |||| ||||| |  |: |:|||:|||::||
Db     99 IRARQSRDYLQRDCPQQLKQVLELQRGVLGQQVPTLVKVTRHWASTGTSLRCQALNFFPQ 158

Qy    212 NITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPL 271
          ||||:|||| ||:|||:  |::|||||||||||:|||| |||| |::||||||||||||
Db    159 NITMRWLKDSQPLDAKDVNPENVLPNGDGTYQGWLTLAVAPGEETRFSCQVEHPGLDQPL 218

Qy    272 IVIWE 276
             ||
Db    219 TATWE 223


RESULT 13
Q9HC83
ID   Q9HC83      PRELIMINARY;      PRT;   161 AA.
AC   Q9HC83;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hemochromatosis splice variant delE3-7.
GN   HFE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20448010; PubMed=11001625;
RA   Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.Y., David V.,
RA   Mosser J.;
RT   "The HFE gene undergoes alternate splicing processes.";
RL   Blood Cells Mol. Dis. 26:155-162(2000).
DR   EMBL; AF115264; AAG29571.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00129; MHC_I; 1.
```

```
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
SQ   SEQUENCE   161 AA;  18651 MW;  5E288C5835DC3784 CRC64;

  Query Match             43.5%;  Score 662;  DB 4;  Length 161;
  Best Local Similarity   100.0%;  Pred. No. 2e-56;
  Matches  121;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 RLLRSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQ 82

Qy      61 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 MWLQLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDG 142

Qy     121 Q 121
           |
Db     143 Q 143


RESULT 14
Q9HC69
ID   Q9HC69      PRELIMINARY;      PRT;   116 AA.
AC   Q9HC69;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hemochromatosis splice variant 861-2305del (Fragment).
GN   HFE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20448010; PubMed=11001625;
RA   Thenie A., Orhant M., Gicquel I., Fergelot P., Le Gall J.Y., David V.,
RA   Mosser J.;
RT   "The HFE gene undergoes alternate splicing processes.";
RL   Blood Cells Mol. Dis. 26:155-162(2000).
DR   EMBL; AF144241; AAG29576.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   116 AA;  13541 MW;  AC0333B096A3F47B CRC64;

  Query Match             38.9%;  Score 592;  DB 4;  Length 116;
  Best Local Similarity   99.1%;  Pred. No. 8e-50;
  Matches  106;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      89 HSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWE 148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 HSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEPRAWPTKLEWE 60

Qy     149 RHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVT 195
           |||||||||||||||||||||||||||||||||||||||||||||||:
Db      61 RHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVS 107


RESULT 15
Q9BD50
ID   Q9BD50      PRELIMINARY;      PRT;   340 AA.
AC   Q9BD50;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   MHC class I related protein MR1 isoform.
GN   MR1.
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Parra-Cuadrado J.F., Garcia-Pavia P., Gomez del Moral M.,
RA   Martinez-Naves E.;
RT   "Identification of MR1 cDNA sequences in non-human primates.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC       IMMUNE SYSTEM (BY SIMILARITY).
CC   -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC       MICROGLOBULIN) (BY SIMILARITY).
DR   EMBL; AJ271828; CAC28215.1; -.
DR   HSSP; Q30201; 1A6Z.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001039; MHC_I.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00129; MHC_I; 1.
DR   PRINTS; PR01638; MHCCLASSI.
DR   ProDom; PD000050; MHC_I; 1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS00290; IG_MHC; 1.
KW   Glycoprotein; Transmembrane.
SQ   SEQUENCE   340 AA;  39375 MW;  A893952B78725F17 CRC64;

  Query Match             35.6%;  Score 542.5;  DB 7;  Length 340;
  Best Local Similarity   39.5%;  Pred. No. 2.1e-44;
  Matches  107;  Conservative   51;  Mismatches  110;  Indels    3;  Gaps    3;

Qy       4 RSHSLHYLFMGASEQDLGLSLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWL 63
           |:||| |  :| |:   |:  | ::|||      ||  ::: ||| ||::  ::   |
Db      23 RTHSLRYFRLGVSDPIRGVPEFISVGYVDSHPITTYDSVTQQKEPRAPWMAENLAPDHWE 82

Qy      64 QLSQSLKGWDHMFTVDFWTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDH 123
           : :| |:||  || |:   :  ::||| ||| | ::|||: || || |: :| |||||
Db      83 RYTQLLRGWQQMFKVELKRLQRHYNHS-GSHTYQRMIGCELLEDGSTTGFLQYAYDGQDF 141

Qy     124 LEFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQ 183
           | |  ||| | | :  |   |  || ::   :  : :|| :| | |:: || |:  | :
Db     142 LIFNKDTLSWLAVDNVAHTIKRAWEANQHELQYQKNWLEEECIAWLKRFLEYGKDTLQRT 201

Qy     184 VPPLVKVTHHVT-SSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTY 242
            ||||:|    |   |||| |:|  :||  | | |:|: : :  :| :  |:||:|||||
Db     202 EPPLVRVNRKETFPGVTTLFCKAHGFYPPEIYMTWMKNGEEI-VQEMDYGDILPSGDGTY 260

Qy     243 QGWITLAVPPGEEQRYTCQVEHPGLDQPLIV 273
           | | :  : |     |:| ||| |:   | |
Db     261 QTWASFELDPQSSNLYSCHVEHCGVHMVLQV 291


Search completed: March 31, 2003, 14:09:50
Job time : 35 secs
```

Connecting via Winsock to STN

Welcome to STN International! Enter x:x

LOGINID:ssspta1644axd

PASSWORD:
TERMINAL (ENTER 1, 2, 3, OR ?):2

* * * * * * * * * * Welcome to STN International * * * * * * * * * *

NEWS 1		Web Page URLs for STN Seminar Schedule - N. America
NEWS 2	Apr 08	"Ask CAS" for self-help around the clock
NEWS 3	Apr 09	BEILSTEIN: Reload and Implementation of a New Subject Area
NEWS 4	Apr 09	ZDB will be removed from STN
NEWS 5	Apr 19	US Patent Applications available in IFICDB, IFIPAT, and IFIUDB
NEWS 6	Apr 22	Records from IP.com available in CAPLUS, HCAPLUS, and ZCAPLUS
NEWS 7	Apr 22	BIOSIS Gene Names now available in TOXCENTER
NEWS 8	Apr 22	Federal Research in Progress (FEDRIP) now available
NEWS 9	Jun 03	New e-mail delivery for search results now available
NEWS 10	Jun 10	MEDLINE Reload
NEWS 11	Jun 10	PCTFULL has been reloaded
NEWS 12	Jul 02	FOREGE no longer contains STANDARDS file segment
NEWS 13	Jul 22	USAN to be reloaded July 28, 2002; saved answer sets no longer valid
NEWS 14	Jul 29	Enhanced polymer searching in REGISTRY
NEWS 15	Jul 30	NETFIRST to be removed from STN
NEWS 16	Aug 08	CANCERLIT reload
NEWS 17	Aug 08	PHARMAMarketLetter(PHARMAML) - new on STN
NEWS 18	Aug 08	NTIS has been reloaded and enhanced
NEWS 19	Aug 19	Aquatic Toxicity Information Retrieval (AQUIRE) now available on STN
NEWS 20	Aug 19	IFIPAT, IFICDB, and IFIUDB have been reloaded
NEWS 21	Aug 19	The MEDLINE file segment of TOXCENTER has been reloaded
NEWS 22	Aug 26	Sequence searching in REGISTRY enhanced
NEWS 23	Sep 03	JAPIO has been reloaded and enhanced
NEWS 24	Sep 16	Experimental properties added to the REGISTRY file
NEWS 25	Sep 16	CA Section Thesaurus available in CAPLUS and CA
NEWS 26	Oct 01	CASREACT Enriched with Reactions from 1907 to 1985
NEWS 27	Oct 21	EVENTLINE has been reloaded
NEWS 28	Oct 24	BEILSTEIN adds new search fields
NEWS 29	Oct 24	Nutraceuticals International (NUTRACEUT) now available on STN
NEWS 30	Oct 25	MEDLINE SDI run of October 8, 2002
NEWS 31	Nov 18	DKILIT has been renamed APOLLIT
NEWS 32	Nov 25	More calculated properties added to REGISTRY
NEWS 33	Dec 02	TIBKAT will be removed from STN
NEWS 34	Dec 04	CSA files on STN
NEWS 35	Dec 17	PCTFULL now covers WP/PCT Applications from 1978 to date
NEWS 36	Dec 17	TOXCENTER enhanced with additional content
NEWS 37	Dec 17	Adis Clinical Trials Insight now available on STN
NEWS 38	Dec 30	ISMEC no longer available
NEWS 39	Jan 21	NUTRACEUT offering one free connect hour in February 2003
NEWS 40	Jan 21	PHARMAML offering one free connect hour in February 2003
NEWS 41	Jan 29	Simultaneous left and right truncation added to COMPENDEX, ENERGY, INSPEC
NEWS 42	Feb 13	CANCERLIT is no longer being updated
NEWS 43	Feb 24	METADEX enhancements
NEWS 44	Feb 24	PCTGEN now available on STN
NEWS 45	Feb 24	TEMA now available on STN

```
 NEWS 46   Feb 26  NTIS now allows simultaneous left and right truncation
 NEWS 47   Feb 26  PCTFULL now contains images
 NEWS 48   Mar 04  SDI PACKAGE for monthly delivery of multifile SDI results
 NEWS 49   Mar 19  APOLLIT offering free connect time in April 2003
 NEWS 50   Mar 20  EVENTLINE will be removed from STN
 NEWS 51   Mar 24  PATDPAFULL now available on STN
 NEWS 52   Mar 24  Additional information for trade-named substances without
                   structures available in REGISTRY
 NEWS 53   Mar 24  Indexing from 1957 to 1966 added to records in CA/CAPLUS

 NEWS EXPRESS  January 6 CURRENT WINDOWS VERSION IS V6.01a,
               CURRENT MACINTOSH VERSION IS V6.0b(ENG) AND V6.0Jb(JP),
               AND CURRENT DISCOVER FILE IS DATED 01 OCTOBER 2002
 NEWS HOURS    STN Operating Hours Plus Help Desk Availability
 NEWS INTER    General Internet Information
 NEWS LOGIN    Welcome Banner and News Items
 NEWS PHONE    Direct Dial and Telecommunication Network Access to STN
 NEWS WWW      CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that
specific topic.

  All use of STN is subject to the provisions of the STN Customer
  agreement.  Please note that this agreement limits use to scientific
  research.  Use for software development or design or implementation
  of commercial gateways or other similar uses is prohibited and may
  result in loss of user privileges and other penalties.

  * * * * * * * * * * * * * STN  Columbus  * * * * * * * * * * * * * *



FILE 'HOME' ENTERED AT 13:37:17 ON 31 MAR 2003

=>  file medline caplus embase biosis
COST IN U.S. DOLLARS                                  SINCE FILE      TOTAL
                                                         ENTRY    SESSION
FULL ESTIMATED COST                                       0.21       0.21

FILE 'MEDLINE' ENTERED AT 13:37:30 ON 31 MAR 2003

FILE 'CAPLUS' ENTERED AT 13:37:30 ON 31 MAR 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'EMBASE' ENTERED AT 13:37:30 ON 31 MAR 2003
COPYRIGHT (C) 2003 Elsevier Science B.V.  All rights reserved.

FILE 'BIOSIS' ENTERED AT 13:37:30 ON 31 MAR 2003
COPYRIGHT (C) 2003 BIOLOGICAL ABSTRACTS INC.(R)

=> s feder J?/au or bjorkman P?/au or Schatzman R?/au
L1         1449 FEDER J?/AU OR BJORKMAN P?/AU OR SCHATZMAN R?/AU

=> s l1 and HFE
L2           91 L1 AND HFE

=> s l1 and PD<19980612
'19980612' NOT A VALID FIELD CODE
   3 FILES SEARCHED...
L3          923 L1 AND PD<19980612
```

=> s l2 and soluble
L4 28 L2 AND SOLUBLE

=> dup rem l4
PROCESSING COMPLETED FOR L4
L5 9 DUP REM L4 (19 DUPLICATES REMOVED)

=> dis l5 1-9 ibib abs kwic

L5 ANSWER 1 OF 9 MEDLINE DUPLICATE 1
ACCESSION NUMBER: 2001520424 MEDLINE
DOCUMENT NUMBER: 21433894 PubMed ID: 11425849
TITLE: Hydrophobic ligand binding by Zn-alpha 2-glycoprotein, a **soluble** fat-depleting factor related to major histocompatibility complex proteins.
AUTHOR: Kennedy M W; Heikema A P; Cooper A; **Bjorkman P J**; Sanchez L M
CORPORATE SOURCE: Division of Environmental and Evolutionary Biology, Institute of Biomedical and Life Sciences and the Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, United Kingdom.. malcolm.kennedy@bio.gla.ac.uk
SOURCE: JOURNAL OF BIOLOGICAL CHEMISTRY, (2001 Sep 14) 276 (37) 35008-13.
Journal code: 2985121R. ISSN: 0021-9258.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200110
ENTRY DATE: Entered STN: 20010925
Last Updated on STN: 20030105
Entered Medline: 20011011

AB Zn-alpha(2)-glycoprotein (ZAG) is a member of the major histocompatibility complex (MHC) class I family of proteins and is identical in amino acid sequence to a tumor-derived lipid-mobilizing factor associated with cachexia in cancer patients. ZAG is present in plasma and other body fluids, and its natural function, like leptin's, probably lies in lipid store homeostasis. X-ray crystallography has revealed an open groove between the helices of ZAG's alpha(1) and alpha(2) domains, containing an unidentified small ligand in a position similar to that of peptides in MHC proteins (Sanchez, L. M., Chirino, A. J., and Bjorkman, P. J. (1999) Science 283, 1914-1919). Here we show, using serum-derived and bacterial recombinant protein, that ZAG binds the fluorophore-tagged fatty acid 11-(dansylamino)undecanoic acid (DAUDA) and, by competition, natural fatty acids such as arachidonic, linolenic, eicosapentaenoic, and docosahexaenoic acids. Other MHC class I-related proteins (FcRn, **HFE**, HLA-Cw*0702) showed no such evidence of binding. Fluorescence and isothermal calorimetry analysis showed that ZAG binds DAUDA with K(d) in the micromolar range, and differential scanning calorimetry showed that ligand binding increases the thermal stability of the protein. Addition of fatty acids to ZAG alters its intrinsic (tryptophan) fluorescence emission spectrum, providing a strong indication that ligand binds in the expected position close to a cluster of exposed tryptophan side chains in the groove. This study therefore shows that ZAG binds small hydrophobic ligands, that the natural ligand may be a polyunsaturated fatty acid, and provides a fluorescence-based method for investigating ZAG-ligand interactions.

TI Hydrophobic ligand binding by Zn-alpha 2-glycoprotein, a **soluble** fat-depleting factor related to major histocompatibility complex proteins.
AU Kennedy M W; Heikema A P; Cooper A; **Bjorkman P J**; Sanchez L M

AB . . . and, by competition, natural fatty acids such as arachidonic, linolenic, eicosapentaenoic, and docosahexaenoic acids. Other MHC class I-related proteins (FcRn, **HFE**, HLA-Cw*0702) showed no such evidence of binding. Fluorescence and isothermal calorimetry analysis

showed that ZAG binds DAUDA with K(d) in. . .

L5 ANSWER 2 OF 9 MEDLINE DUPLICATE 2
ACCESSION NUMBER: 2002073391 MEDLINE
DOCUMENT NUMBER: 21659940 PubMed ID: 11800564
TITLE: Mutational analysis of the transferrin receptor reveals overlapping **HFE** and transferrin binding sites.
AUTHOR: West A P Jr; Giannetti A M; Herr A B; Bennett M J; Nangiana J S; Pierce J R; Weiner L P; Snow P M; **Bjorkman P J**
CORPORATE SOURCE: Division of Biology 156-29 , California Institute of Technology, Pasadena, CA 91125, USA.
CONTRACT NUMBER: 5T32 GM07616 (NIGMS)
SOURCE: JOURNAL OF MOLECULAR BIOLOGY, (2001 Oct 19) 313 (2) 385-97. Journal code: 2985088R. ISSN: 0022-2836.
PUB. COUNTRY: England: United Kingdom
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200202
ENTRY DATE: Entered STN: 20020125
Last Updated on STN: 20020205
Entered Medline: 20020204

AB The transferrin receptor (TfR) binds two proteins critical for iron metabolism: transferrin (Tf) and **HFE**, the protein mutated in hereditary hemochromatosis. Previous results demonstrated that Tf and **HFE** compete for binding to TfR, suggesting that Tf and **HFE** bind to the same or an overlapping site on TfR. TfR is a homodimer that binds one Tf per polypeptide chain (2:2, TfR/Tf stoichiometry), whereas both 2:1 and 2:2 TfR/**HFE** stoichiometries have been observed. In order to more fully characterize the interaction between **HFE** and TfR, we determined the binding stoichiometry using equilibrium gel-filtration and analytical ultracentrifugation. Both techniques indicate that a 2:2 TfR/**HFE** complex can form at submicromolar concentrations in solution, consistent with the hypothesis that **HFE** competes for Tf binding to TfR by blocking the Tf binding site rather than by exerting an allosteric effect. To determine whether the Tf and **HFE** binding sites on TfR overlap, residues at the **HFE** binding site on TfR were identified from the 2.8 A resolution **HFE**-TfR co-crystal structure, then mutated and tested for their effects on **HFE** and Tf binding. The binding affinities of **soluble** TfR mutants for **HFE** and Tf were determined using a surface plasmon resonance assay. Substitutions of five TfR residues at the **HFE** binding site (L619A, R629A, Y643A, G647A and F650Q) resulted in significant reductions in Tf binding affinity. The findings that both **HFE** and Tf form 2:2 complexes with TfR and that mutations at the **HFE** binding site affect Tf binding support a model in which **HFE** and Tf compete for overlapping binding sites on TfR.

TI Mutational analysis of the transferrin receptor reveals overlapping **HFE** and transferrin binding sites.
AU. . . A M; Herr A B; Bennett M J; Nangiana J S; Pierce J R; Weiner L P; Snow P M; **Bjorkman P J**
AB The transferrin receptor (TfR) binds two proteins critical for iron metabolism: transferrin (Tf) and **HFE**, the protein mutated in hereditary hemochromatosis. Previous results demonstrated that Tf and **HFE** compete for binding to TfR, suggesting that Tf and **HFE** bind to the same or an overlapping site on TfR. TfR is a homodimer that binds one Tf per polypeptide chain (2:2, TfR/Tf stoichiometry), whereas both 2:1 and 2:2 TfR/**HFE** stoichiometries have been observed. In order to more fully characterize the interaction between **HFE** and TfR, we determined the binding stoichiometry using equilibrium gel-filtration and analytical ultracentrifugation. Both techniques indicate that a 2:2 TfR/**HFE** complex can form at submicromolar

concentrations in solution, consistent with the hypothesis that **HFE** competes for Tf binding to TfR by blocking the Tf binding site rather than by exerting an allosteric effect. To determine whether the Tf and **HFE** binding sites on TfR overlap, residues at the **HFE** binding site on TfR were identified from the 2.8 A resolution **HFE**-TfR co-crystal structure, then mutated and tested for their effects on **HFE** and Tf binding. The binding affinities of **soluble** TfR mutants for **HFE** and Tf were determined using a surface plasmon resonance assay. Substitutions of five TfR residues at the **HFE** binding site (L619A, R629A, Y643A, G647A and F650Q) resulted in significant reductions in Tf binding affinity. The findings that both **HFE** and Tf form 2:2 complexes with TfR and that mutations at the **HFE** binding site affect Tf binding support a model in which **HFE** and Tf compete for overlapping binding sites on TfR.

L5 ANSWER 3 OF 9 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 2001:250785 BIOSIS
DOCUMENT NUMBER: PREV200100250785
TITLE: Interactions of the ectodomain of **HFE** with the transferrin receptor are critical for iron homeostasis in cells.
AUTHOR(S): Enns, Caroline A. (1); Roy, Cindy N. (1); **Feder, John N.**
CORPORATE SOURCE: (1) Oregon Health Sciences University, 3181 SW Sam Jackson Park Rd., Portand, OR, 97201 USA
SOURCE: FASEB Journal, (March 7, 2001) Vol. 15, No. 4, pp. A60. print.
Meeting Info.: Annual Meeting of the Federation of American Societies for Experimental Biology on Experimental Biology 2001 Orlando, Florida, USA March 31-April 04, 2001
ISSN: 0892-6638.
DOCUMENT TYPE: Conference
LANGUAGE: English
SUMMARY LANGUAGE: English
AB Expression of wild type **HFE** reduces the ferritin levels of cells in culture. In this report we demonstrate that the predominant hereditary hemochromatosis mutation, C282Y **HFE**, does not reduce ferritin expression. However, the second mutation, H63D **HFE**, reduces ferritin expression to a level indistinguishable from cells expressing wild type **HFE**. Further, two mutations in the cytoplasmic domain of **HFE** engineered to disrupt potential phosphorylation-mediated signal transduction events, S335M and Y342C, were functionally indistinguishable from wild type **HFE** in this assay, as was **soluble HFE**. These results implicate a role for the interaction of **HFE** with the transferrin receptor in lowering cellular ferritin levels.
TI Interactions of the ectodomain of **HFE** with the transferrin receptor are critical for iron homeostasis in cells.
AU Enns, Caroline A. (1); Roy, Cindy N. (1); **Feder, John N.**
AB Expression of wild type **HFE** reduces the ferritin levels of cells in culture. In this report we demonstrate that the predominant hereditary hemochromatosis mutation, C282Y **HFE**, does not reduce ferritin expression. However, the second mutation, H63D **HFE**, reduces ferritin expression to a level indistinguishable from cells expressing wild type **HFE**. Further, two mutations in the cytoplasmic domain of **HFE** engineered to disrupt potential phosphorylation-mediated signal transduction events, S335M and Y342C, were functionally indistinguishable from wild type **HFE** in this assay, as was **soluble HFE**. These results implicate a role for the interaction of **HFE** with the transferrin receptor in lowering cellular ferritin levels.
IT Major Concepts

Molecular Genetics (Biochemistry and Molecular Biophysics); Metabolism
IT Chemicals & Biochemicals
HFE: ectodomain; iron: homeostasis; transferrin receptor
GEN **HFE** gene: mutation

L5 ANSWER 4 OF 9 MEDLINE DUPLICATE 3
ACCESSION NUMBER: 2001088821 MEDLINE
DOCUMENT NUMBER: 20556244 PubMed ID: 11027676
TITLE: Comparison of the interactions of transferrin receptor and transferrin receptor 2 with transferrin and the hereditary hemochromatosis protein **HFE**.
AUTHOR: West A P Jr; Bennett M J; Sellers V M; Andrews N C; Enns C A; **Bjorkman P J**
CORPORATE SOURCE: Division of Biology and Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California 91125, USA.
CONTRACT NUMBER: DK54488 (NIDDK)
SOURCE: JOURNAL OF BIOLOGICAL CHEMISTRY, (2000 Dec 8) 275 (49) 38135-8.
Journal code: 2985121R. ISSN: 0021-9258.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200101
ENTRY DATE: Entered STN: 20010322
Last Updated on STN: 20010322
Entered Medline: 20010118

AB The transferrin receptor (TfR) interacts with two proteins important for iron metabolism, transferrin (Tf) and **HFE**, the protein mutated in hereditary hemochromatosis. A second receptor for Tf, TfR2, was recently identified and found to be functional for iron uptake in transfected cells (Kawabata, H., Germain, R. S., Vuong, P. T., Nakamaki, T., Said, J. W., and Koeffler, H. P. (2000) J. Biol. Chem. 275, 16618-16625). TfR2 has a pattern of expression and regulation that is distinct from TfR, and mutations in TfR2 have been recognized as the cause of a non-**HFE** linked form of hemochromatosis (Camaschella, C., Roetto, A., Cali, A., De Gobbi, M., Garozzo, G., Carella, M., Majorano, N., Totaro, A., and Gasparini, P. (2000) Nat. Genet. 25, 14-15). To investigate the relationship between TfR, TfR2, Tf, and **HFE**, we performed a series of binding experiments using **soluble** forms of these proteins. We find no detectable binding between TfR2 and **HFE** by co-immunoprecipitation or using a surface plasmon resonance-based assay. The affinity of TfR2 for iron-loaded Tf was determined to be 27 nm, 25-fold lower than the affinity of TfR for Tf. These results imply that **HFE** regulates Tf-mediated iron uptake only from the classical TfR and that TfR2 does not compete for **HFE** binding in cells expressing both forms of TfR.

TI Comparison of the interactions of transferrin receptor and transferrin receptor 2 with transferrin and the hereditary hemochromatosis protein **HFE**.

AU West A P Jr; Bennett M J; Sellers V M; Andrews N C; Enns C A; **Bjorkman P J**

AB The transferrin receptor (TfR) interacts with two proteins important for iron metabolism, transferrin (Tf) and **HFE**, the protein mutated in hereditary hemochromatosis. A second receptor for Tf, TfR2, was recently identified and found to be functional. . . expression and regulation that is distinct from TfR, and mutations in TfR2 have been recognized as the cause of a non-**HFE** linked form of hemochromatosis (Camaschella, C., Roetto, A., Cali, A., De Gobbi, M., Garozzo, G., Carella, M., Majorano, N., Totaro, A., and Gasparini, P. (2000) Nat. Genet. 25, 14-15). To investigate the relationship between TfR, TfR2, Tf, and **HFE**, we performed a series of binding experiments using **soluble** forms of these proteins. We find no

detectable binding between TfR2 and **HFE** by co-immunoprecipitation or using a surface plasmon resonance-based assay. The affinity of TfR2 for iron-loaded Tf was determined to be 27 nm, 25-fold lower than the affinity of TfR for Tf. These results imply that **HFE** regulates Tf-mediated iron uptake only from the classical TfR and that TfR2 does not compete for **HFE** binding in cells expressing both forms of TfR.

L5 ANSWER 5 OF 9 MEDLINE DUPLICATE 4
ACCESSION NUMBER: 2001061091 MEDLINE
DOCUMENT NUMBER: 20532540 PubMed ID: 11078891
TITLE: Interactions of the ectodomain of **HFE** with the transferrin receptor are critical for iron homeostasis in cells.
AUTHOR: Roy C N; Carlson E J; Anderson E L; Basava A; Starnes S M; **Feder J N**; Enns C A
CORPORATE SOURCE: Department of Cell and Developmental Biology, Oregon Health Sciences University, Portland 97201-3098, USA.
CONTRACT NUMBER: DK 54488 (NIDDK)
T32-HL00781 (NHLBI)
SOURCE: FEBS LETTERS, (2000 Nov 10) 484 (3) 271-4.
Journal code: 0155157. ISSN: 0014-5793.
PUB. COUNTRY: Netherlands
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200012
ENTRY DATE: Entered STN: 20010322
Last Updated on STN: 20010322
Entered Medline: 20001222
AB Expression of wild type **HFE** reduces the ferritin levels of cells in culture. In this report we demonstrate that the predominant hereditary hemochromatosis mutation, C282Y(2) **HFE**, does not reduce ferritin expression. However, the second mutation, H63D **HFE**, reduces ferritin expression to a level indistinguishable from cells expressing wild type **HFE**. Further, two **HFE** cytoplasmic domain mutations engineered to disrupt potential signal transduction, S335M and Y342C, were functionally indistinguishable from wild type **HFE** in this assay, as was **soluble HFE**. These results implicate a role for the interaction of **HFE** with the transferrin receptor in lowering cellular ferritin levels.
TI Interactions of the ectodomain of **HFE** with the transferrin receptor are critical for iron homeostasis in cells.
AU Roy C N; Carlson E J; Anderson E L; Basava A; Starnes S M; **Feder J N**; Enns C A
AB Expression of wild type **HFE** reduces the ferritin levels of cells in culture. In this report we demonstrate that the predominant hereditary hemochromatosis mutation, C282Y(2) **HFE**, does not reduce ferritin expression. However, the second mutation, H63D **HFE**, reduces ferritin expression to a level indistinguishable from cells expressing wild type **HFE**. Further, two **HFE** cytoplasmic domain mutations engineered to disrupt potential signal transduction, S335M and Y342C, were functionally indistinguishable from wild type **HFE** in this assay, as was **soluble HFE**. These results implicate a role for the interaction of **HFE** with the transferrin receptor in lowering cellular ferritin levels.

L5 ANSWER 6 OF 9 MEDLINE DUPLICATE 5
ACCESSION NUMBER: 2000027103 MEDLINE
DOCUMENT NUMBER: 20027103 PubMed ID: 10556042
TITLE: The hemochromatosis protein **HFE** competes with transferrin for binding to the transferrin receptor.
AUTHOR: Lebron J A; West A P Jr; **Bjorkman P J**
CORPORATE SOURCE: Division of Biology, California Institute of Technology,

1200 East California Boulevard, Pasadena, CA 91125, USA.
CONTRACT NUMBER: GM07616 (NIGMS)
SOURCE: JOURNAL OF MOLECULAR BIOLOGY, (1999 Nov 19) 294 (1) 239-45.
Journal code: 2985088R. ISSN: 0022-2836.
PUB. COUNTRY: ENGLAND: United Kingdom
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 200001
ENTRY DATE: Entered STN: 20000124
Last Updated on STN: 20000124
Entered Medline: 20000110

AB **HFE** is a class I major histocompatibility complex (MHC)-related protein that is mutated in patients with the iron overload disease hereditary hemochromatosis. **HFE** binds to transferrin receptor (TfR), the receptor used by cells to obtain iron in the form of diferric transferrin (Fe-Tf). Previous studies demonstrated that **HFE** and Fe-Tf can bind simultaneously to TfR to form a ternary complex, and that membrane-bound or **soluble HFE** binding to cell surface TfR results in a reduction in the affinity of TfR for Fe-Tf. We studied the inhibition by **soluble HFE** of the interaction between **soluble** TfR and Fe-Tf using radioactivity-based and biosensor-based assays. The results demonstrate that **HFE** inhibits the TfR:Fe-Tf interaction by binding at or near the Fe-Tf binding site on TfR, and that the Fe-Tf:TfR:**HFE** ternary complex consists of one Fe-Tf and one **HFE** bound to a TfR homodimer.
Copyright 1999 Academic Press.

TI The hemochromatosis protein **HFE** competes with transferrin for binding to the transferrin receptor.

AU Lebron J A; West A P Jr; **Bjorkman P J**

AB **HFE** is a class I major histocompatibility complex (MHC)-related protein that is mutated in patients with the iron overload disease hereditary hemochromatosis. **HFE** binds to transferrin receptor (TfR), the receptor used by cells to obtain iron in the form of diferric transferrin (Fe-Tf). Previous studies demonstrated that **HFE** and Fe-Tf can bind simultaneously to TfR to form a ternary complex, and that membrane-bound or **soluble HFE** binding to cell surface TfR results in a reduction in the affinity of TfR for Fe-Tf. We studied the inhibition by **soluble HFE** of the interaction between **soluble** TfR and Fe-Tf using radioactivity-based and biosensor-based assays. The results demonstrate that **HFE** inhibits the TfR:Fe-Tf interaction by binding at or near the Fe-Tf binding site on TfR, and that the Fe-Tf:TfR:**HFE** ternary complex consists of one Fe-Tf and one **HFE** bound to a TfR homodimer.
Copyright 1999 Academic Press.

L5 ANSWER 7 OF 9 MEDLINE DUPLICATE 6
ACCESSION NUMBER: 1998132614 MEDLINE
DOCUMENT NUMBER: 98132614 PubMed ID: 9465039
TITLE: The hemochromatosis gene product complexes with the transferrin receptor and lowers its affinity for ligand binding.
AUTHOR: **Feder J N**; Penny D M; Irrinki A; Lee V K; Lebron J A; Watson N; Tsuchihashi Z; Sigal E; **Bjorkman P J**; **Schatzman R C**
CORPORATE SOURCE: Progenitor, Inc. (formerly Mercator Genetics, Inc.), 4040 Campbell Avenue, Menlo Park, CA 94025, USA..
feder@progenitor.com
SOURCE: PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1998 Feb 17) 95 (4) 1472-7.
Journal code: 7505876. ISSN: 0027-8424.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English

FILE SEGMENT: Priority Journals
ENTRY MONTH: 199803
ENTRY DATE: Entered STN: 19980326
Last Updated on STN: 19980326
Entered Medline: 19980319

AB We recently reported the positional cloning of a candidate gene for hereditary hemochromatosis called **HFE**. The gene product, a member of the major histocompatibility complex class I-like family, was found to have a mutation, Cys-282 --> Tyr (C282Y), in 85% of patient chromosomes. This mutation eliminates the ability of **HFE** to associate with beta2-microglobulin (beta2m) and prevents cell-surface expression. A second mutation that has no effect on beta2m association, H63D, was found in eight out of nine patients heterozygous for the C282Y mutant. In this report, we demonstrate in cultured 293 cells overexpressing wild-type or mutant **HFE** proteins that both the wild-type and H63D **HFE** proteins form stable complexes with the transferrin receptor (TfR). The C282Y mutation nearly completely prevents the association of the mutant **HFE** protein with the TfR. Studies on cell-associated transferrin at 37 degrees C suggest that the overexpressed wild-type **HFE** protein decreases the affinity of the TfR for transferrin. The overexpressed H63D protein does not have this effect, providing the first direct evidence for a functional consequence of the H63D mutation. Addition of **soluble** wild-type **HFE** /beta2m heterodimers to cultured cells also decreased the apparent affinity of the TfR for its ligand under steady-state conditions, both in 293 cells and in HeLa cells. Furthermore, at 4 degrees C, the added **soluble** complex of **HFE**/beta2m inhibited binding of transferrin to HeLa cell TfR in a concentration-dependent manner. Scatchard plots of these data indicate that the added heterodimer substantially reduced the affinity of TfR for transferrin. These results establish a molecular link between **HFE** and a key protein involved in iron transport, the TfR, and raise the possibility that alterations in this regulatory mechanism may play a role in the pathogenesis of hereditary hemochromatosis.

AU **Feder J N**; Penny D M; Irrinki A; Lee V K; Lebron J A; Watson N; Tsuchihashi Z; Sigal E; **Bjorkman P J**; **Schatzman R C**

AB We recently reported the positional cloning of a candidate gene for hereditary hemochromatosis called **HFE**. The gene product, a member of the major histocompatibility complex class I-like family, was found to have a mutation, Cys-282 --> Tyr (C282Y), in 85% of patient chromosomes. This mutation eliminates the ability of **HFE** to associate with beta2-microglobulin (beta2m) and prevents cell-surface expression. A second mutation that has no effect on beta2m association, H63D,. . . nine patients heterozygous for the C282Y mutant. In this report, we demonstrate in cultured 293 cells overexpressing wild-type or mutant **HFE** proteins that both the wild-type and H63D **HFE** proteins form stable complexes with the transferrin receptor (TfR). The C282Y mutation nearly completely prevents the association of the mutant **HFE** protein with the TfR. Studies on cell-associated transferrin at 37 degrees C suggest that the overexpressed wild-type **HFE** protein decreases the affinity of the TfR for transferrin. The overexpressed H63D protein does not have this effect, providing the first direct evidence for a functional consequence of the H63D mutation. Addition of **soluble** wild-type **HFE**/beta2m heterodimers to cultured cells also decreased the apparent affinity of the TfR for its ligand under steady-state conditions, both in 293 cells and in HeLa cells. Furthermore, at 4 degrees C, the added **soluble** complex of **HFE**/beta2m inhibited binding of transferrin to HeLa cell TfR in a concentration-dependent manner. Scatchard plots of these data indicate that the added heterodimer substantially reduced the affinity of TfR for transferrin. These results establish a molecular link between **HFE** and a key protein involved in iron transport, the TfR, and raise the possibility that alterations in this regulatory mechanism. . .

L5 ANSWER 8 OF 9 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 1998:526310 BIOSIS
DOCUMENT NUMBER: PREV199800526310
TITLE: The hemochromatosis gene product (**HFE**) is present in **soluble** form in serum and is related to body iron content.
AUTHOR(S): Li, Q. (1); Hammett, R. J. H. (1); Battaglia, E. (1); **Feder, J. N.**; Gollan, J. L. (1)
CORPORATE SOURCE: (1) Brigham and Women's Hosp., Harvard Med. Sch., Boston, MA USA
SOURCE: Hepatology, (Oct., 1998) Vol. 28, No. 4 PART 2, pp. 501A. Meeting Info.: Biennial Scientific Meeting of the International Association for the Study of the Liver and the 49th Annual Meeting and Postgraduate Courses of the American Association for the Study of Liver Diseases Chicago, Illinois, USA November 4-10, 1998 International Association for the Study of the Liver . ISSN: 0270-9139.
DOCUMENT TYPE: Conference
LANGUAGE: English
TI The hemochromatosis gene product (**HFE**) is present in **soluble** form in serum and is related to body iron content.
AU Li, Q. (1); Hammett, R. J. H. (1); Battaglia, E. (1); **Feder, J. N.** ; Gollan, J. L. (1)

L5 ANSWER 9 OF 9 MEDLINE DUPLICATE 7
ACCESSION NUMBER: 1998206473 MEDLINE
DOCUMENT NUMBER: 98206473 PubMed ID: 9546397
TITLE: Crystal structure of the hemochromatosis protein **HFE** and characterization of its interaction with transferrin receptor.
AUTHOR: Lebron J A; Bennett M J; Vaughn D E; Chirino A J; Snow P M; Mintier G A; **Feder J N**; **Bjorkman P J**
CORPORATE SOURCE: Division of Biology, California Institute of Technology, Pasadena 91125, USA.
SOURCE: CELL, (1998 Apr 3) 93 (1) 111-23. Journal code: 0413066. ISSN: 0092-8674.
PUB. COUNTRY: United States
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199805
ENTRY DATE: Entered STN: 19980514
Last Updated on STN: 19980514
Entered Medline: 19980501
AB **HFE** is an MHC-related protein that is mutated in the iron-overload disease hereditary hemochromatosis. **HFE** binds to transferrin receptor (TfR) and reduces its affinity for iron-loaded transferrin, implicating **HFE** in iron metabolism. The 2.6 A crystal structure of **HFE** reveals the locations of hemochromatosis mutations and a patch of histidines that could be involved in pH-dependent interactions. We also demonstrate that **soluble** TfR and **HFE** bind tightly at the basic pH of the cell surface, but not at the acidic pH of intracellular vesicles. TfR:**HFE** stoichiometry (2:1) differs from TfR:transferrin stoichiometry (2:2), implying a different mode of binding for **HFE** and transferrin to TfR, consistent with our demonstration that **HFE**, transferrin, and TfR form a ternary complex.
TI Crystal structure of the hemochromatosis protein **HFE** and characterization of its interaction with transferrin receptor.
AU Lebron J A; Bennett M J; Vaughn D E; Chirino A J; Snow P M; Mintier G A; **Feder J N**; **Bjorkman P J**
AB **HFE** is an MHC-related protein that is mutated in the iron-overload disease hereditary hemochromatosis. **HFE** binds to

transferrin receptor (TfR) and reduces its affinity for iron-loaded transferrin, implicating **HFE** in iron metabolism. The 2.6 A crystal structure of **HFE** reveals the locations of hemochromatosis mutations and a patch of histidines that could be involved in pH-dependent interactions. We also demonstrate that **soluble** TfR and **HFE** bind tightly at the basic pH of the cell surface, but not at the acidic pH of intracellular vesicles. TfR:**HFE** stoichiometry (2:1) differs from TfR:transferrin stoichiometry (2:2), implying a different mode of binding for **HFE** and transferrin to TfR, consistent with our demonstration that **HFE**, transferrin, and TfR form a ternary complex.

Creation date: 02-27-2004
Indexing Officer: BWYCHE - BRANDON WYCHE
Team: OIPEBackFileIndexing
Dossier: 10092404

Legal Date: 08-05-2003

No.	Doccode	Number of pages
1	SRNT	108

Total number of pages: 108

Remarks:

Order of re-scan issued on .....